The Use of Patient Reported Outcome Measures (PROMs) for Evaluating Emergency Admissions

ESTHER LAAM SUM KWONG

Thesis submitted in accordance with the requirements for the degree of

Doctor of Philosophy of the

University of London

January 2019

Department of Health Services Research and Policy

Faculty of Public Health and Policy

LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE

Funded by the Economic and Social Research Council

Research group affiliation(s):
The research was supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care North Thames

Declaration

I, Esther Laam Sum Kwong, confirm that the work presented in this thesis is my own.

Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature 

Date: 9th January 2019

Abstract

The use of Patient Reported Outcome Measures (PROMs) is one way to measure the effectiveness of health services. PROMs use in emergency admissions to hospitals is limited by the methodological obstacle of having no pre-admission measure of health status. The aim of this thesis was to study the use of retrospective PROMs and of routine measures of population health status to identify a reliable method that would allow the extension of PROMs collection into this important area of health care.

A literature review found there was strong agreement (ICC>0.75) between contemporaneous and retrospective PROMs in elective conditions and that population data could be used to estimate baseline health status in some conditions. This was confirmed in the prospective cohort studies conducted in elective patients (ICC>0.82 for disease-specific and ICC>0.62 for generic PROMs). However, matching methods used to explore population values as an alternative to retrospective baseline health status did not provide estimates similar to those obtained from elective patients. These exploratory matching methods did, however, provide insights for further research in emergency admissions.

The studies in emergency laparotomy (EL) and STEMI heart attack established the feasibility of collecting retrospective PROMs in these emergency settings: 85% of admissions were eligible, and invitation and participation rates were 85% & 72% for EL and 79% & 91% for STEMI. The response rates to three month follow-up questionnaires were good (EL: 74%, STEMI: 66%) enabling the mean recovery of patients in terms of PROMs to be conducted and the effect of any response biases to be determined.

The use of retrospective PROMs can provide a reliable method to collect pre-admission health status in some emergency admissions. It is necessary to establish the generalisability of these findings, investigate possible clinical confounders and explore the extent of unwarranted variation between providers in outcomes.

Acknowledgements

I would like to thank Professor Nick Black my supervisor for his support and encouragement. I deeply value and am extremely grateful for his mentoring and nurturing approach in his supervisions, which continued to inspire me in my curiosity and interest, whilst providing guidance in balancing the need on staying focused throughout. I am grateful to Jenny Neuberger my advisor who became my second supervisor for her patience and considered advice for my statistical and methodological queries.

To both my supervisors, I am thankful for their generosity in providing me timely support throughout to make this thesis possible. I would also like to thank the members of the CLAHRC North Thames Methodological Theme group, for the opportunity to discuss relevant research issues, and for their academic input and constructive criticism on this work at the theme meetings. Finally, and not least, I would like to thank the very special persons who have been there by my side through the ebb and flow, who gave me hope and strength to enjoy and persevere in each step of the way.

Thesis Contents Page

Contents

List of Tables

List of Figures

Abbreviations

AMI	Acute Myocardial Infarction
EL	Emergency Laparotomy
GIQLI	Gastro-Intestinal Quality of Life Index
GP	General Practice
GPPS	General Practice Patient Survey
MINAP	Myocardial Ischaemia National Audit Project
NCAs	National Clinical Audits
NELA	National Emergency Laparotomy Audit
NHS	National Health Service
OHS	Oxford Hip Score
OKS	Oxford Knee Score
PROMs	Patient Reported Outcome Measures
QALYs	Quality Adjust Life Years

QoL	Quality of Life
SAQ-7	Short Version of the Seattle Angina Questionnaire
STEMI	ST Segment Elevation Myocardial Infarction

Chapter 1: Introduction

1.1 Background

1.1.1 Patient reported outcome measures (PROMs)

PROMs are patient self-reported questionnaires that serve to capture a patient's own assessment of their health. They are multi-dimensional measurements of symptoms, functional status, or health-related quality of life. Questionnaires can be used at specific points in time to capture a health change, which can in turn provide the basis for evaluation of an intervention or treatment. PROMs have the potential to change healthcare delivery through assessing relative clinical quality, comparing providers' performance and evaluating the effectiveness of treatments from the perspective of patients. For these reasons, the development of routinely collected PROMs data in four elective surgical procedures in England has been heralded as one of the missing components of quality in the jigsaw in the evaluation of our health service [1,2].

There is growing acceptance of the importance of patients' views of their outcome, in addition to clinical measures such as mortality and morbidity, such that when evaluating interventions and assessing the quality of services, it is necessary to devise ways in which accurate PROMs can be obtained. Development work to widen the use of PROMs also enables the health service to focus on patient-centred care [3].

There is sustained clinical and political interest in further developments of using quality of life measures and outcome focused indicators to evaluate the quality of our health services [4–6]. Currently, PROMs are routinely collected for four elective surgical procedures; these are hip replacement, knee replacement, hernia repair and varicose vein surgeries. Questionnaires are administered prior to the elective surgery during the pre-assessment phase or on admission to quantify a baseline measure of patient self-reported health status (at the current time). Then at a defined time point after the intervention (e.g. 6 months after hip and knee operations), a second questionnaire is administered to re-evaluate the patients' self-reported health status. The

difference between the two questionnaires is taken to be the impact of the intervention. This is usually labelled as health gain, although health status can deteriorate as well as improve [1,7].

Generic PROMs, such as the EQ-5D and SF-36, provide the means to compare health status of patients with different conditions and undergoing different treatments. The generic PROM currently used by NHS England is the EQ-5D score. It asks the patient to score five domains in mobility, self-care, usual activities, pain/discomfort and anxiety/ depression. The measurement is then converted into a utility score using the 'social value' of living in that particular health state based on population valuations derived from a general population survey [8]. This utility score allows for the development of Quality-Adjusted Life Years (QALYs) that are routinely used to inform decision-making in the NHS, including technology appraisals of new medicines by NICE, and economic appraisals [8–10].

Condition-specific PROMs have greater sensitivity by incorporating relevant clinical details specific to a disease. They are more focused on measuring a particular aspect of health which may be of particular use to clinicians and service providers [9]. Examples of condition-specific PROMs include the Oxford Hip Score (OHS) and Oxford Knee Score (OKS) used for hip and knee replacement surgery in the National PROMs Programme in England. Both contain 12 questions on symptoms and activities of daily living validated for each respective condition. For example, the OHS has been developed and validated specifically to assess function and pain for patients undergoing total hip replacement (THR) surgery. The OHS is the most evaluated hip specific measure available [11]. An example question asks "During the past 4 weeks, how would you describe the pain you usually have in your (right/left) (hip/knee)?" Condition-specific PROMs provide greater sensitivity on the condition and can be used to assist shared clinical decision-making. There is emerging evidence of this, for example, joint registries are now incorporating PROMs data into comprehensive benefit-risk assessment tools to support clinical decisions for arthroplasty surgery [12].

Health services and emergency admissions

The NHS has faced an unprecedented period of restricted financial growth in the past decade [13]. There has never been a time when it is more pressing for providers and commissioners to better understand the productivity and quality of services delivered [14]. Measuring the quality of health services can be achieved through assessment of its elements of effectiveness, safety, humanity and equity [15]. Clinical effectiveness can be measured by its components of inputs, processes, and outcomes. The NHS has made significant improvements in its measurement capability in the last two decades through the expansion of robust collection of activity and process data such as hospital episode statistics (primarily for the purposes of reimbursement and payments), alongside clinical measures in the form of national clinical audits for benchmarking and quality improvement. Our health service has faced both external and internal pressures to improve productivity and understanding quality is central to this [14,16]. We must ensure how we measure quality continues to be robust, and relevant to today's healthcare[15].

Since 2008, the Department of Health white paper High Quality care for all had in its strategy to focus specifically on patient centred care, and use patient reported outcomes as a quality measure [17]. This drive has continued through the 2012 health reforms and continues to be the goal for NHS England [6,16]. As survival increases following advances in medical care, mortality rates have generally decreased following hospital admissions and treatment. Therefore, the impact of healthcare on wellbeing and health-related quality of life as a measurable outcome has gained importance. Measuring patient reported outcomes is one of the key ways of capturing this key aspect of quality [2,19]. Indeed, we have seen examples from the routine use of PROMs in elective care that outcomes have a role beyond informing the effectiveness of healthcare, capturing other dimensions of quality such as improving equity of health services through highlighting inequalities in access and provision (e.g. highlighting variations in thresholds for treatment and in outcomes by ethnicity) [20].

Another potential use of routine PROMs, as described earlier, is through determination of Quality Adjusted Life Years (QALYs). This enables economic analysis and cost effectiveness assessments of different treatments in routine clinical practice (rather than in research studies) [9]. However, at present there remains a paucity of outcome data for most routine healthcare within the NHS apart from a small number of elective operations. A review published by York determined the marginal cost to be at £13,000 per QALY for the English NHS, however the authors and others have pointed out that this heavily depended on mortality data as quality of life data currently remains scarce for the majority of health services [21,22]. Therefore, as questions about costs and the quality of care delivered in terms of outcomes matters now more than ever. One of the urgent priorities for the NHS is how to extend the use of PROMs in key areas of the health service.

Finally, PROMs can be used to assess of the diffusion of new medical technologies, novel treatments or the effect of new models of care [23], allowing the impact of such innovations from the patient's perspective to be determined.

The number of emergency admissions has risen by 42% in the 12 years since 2006 [24]. This steep increase in demand has resulted in mismatches of resource allocation within the NHS, demonstrated by the well-documented pressures in our acute hospitals. The nature of emergency admissions has also been changing, with more complex patients being admitted, treated and discharged. Through process measures, we know that the NHS has initially been able to absorb much of this increased demand by decreasing length of stays within hospitals [24,25]. We also know that there remains large variations in demand and processes of care, and survival across the country for emergency admissions [24,26]. What we do not yet know is whether there are similar variations in patient reported outcomes, and how these compare with survival and processes of care.

Although emergency admissions can be considered primarily as lifesaving interventions, as survival rates following emergency acute intervention and hospital care improve, health-related quality of life following the episode

becomes an important outcome. The purpose of interventions in hospital now is not only to save life but to restore patients' health to their full potential. This includes not only their functional capabilities, but also their global wellbeing and quality of life [27]. This outcome of the patients' health following survival also depends greatly on the quality of emergency hospital services, and the care received during their inpatient episode.

Furthermore, clinicians have to exercise some degree of clinical judgement in most cases of emergency treatment, even for life-saving emergency situations and interventions. For example, clinicians have to prioritise the urgency and the timing of conducting the emergency interventions (e.g. emergency laparotomies are classified into categories of urgency e.g. immediate, urgent, and expedited), and make the decision to offer treatment versus offering palliative care. Much of this is based on current clinical evidence of risk-benefit balance to patients, the understanding of the likelihood of survival and likely short-term clinical outcomes. PROMs could offer added understanding of longer-term health status outcomes in these patient groups, and could assist these pertinent clinical decisions.

Without knowing what PROMs are for these emergency conditions, we would never be able to fully understand the effectiveness of the health service in managing patients that require emergency care, and our ability to innovate and improve will be impaired.

1.1.2 The methodological challenge of using PROMs in emergency admissions and possible approaches

However therein lies a methodological challenge as the collection of PROMs can only be completed after, and not before the unexpected, emergency intervention [2]. The challenge for emergency admissions is how to quantify the baseline measurement of patient reported health status, i.e. the prior health status of the patient before the sudden and unexpected onset of ill-health leading to that emergency hospital admission. As it is not feasible to collect a pre-intervention PROM, another method needs to be employed to determine the pre-admission health status.

There are two possible approaches. First is the use of recall, where the patient is asked retrospectively to recollect their health status and health-related quality of life prior to the event. If retrospective reporting were able to provide accurate and reliable measures of previous health status, then the challenge is whether collecting retrospective PROMs following emergency admission is administratively feasible and practical.

An alternative to retrospective questionnaires is to use population values from general population surveys. This is existing data collected routinely on self-reported health status (population values) from surveys such as the UK's General Practice Patient Survey (GPPS) [28].

GPPS is a questionnaire mailed twice a year to approximately 2.7 million adults in total who are registered with a GP in England. Since 2011, the EQ-5D has been incorporated into this survey. There is a response rate of 38%, with nearly 1 million surveys returned annually. A range of patient characteristics are also collected from respondents. Therefore age, sex, and socioeconomic status standardised population EQ-5D values from this source can potentially be used as a proxy for baseline health status measure in place of the retrospective PROM for emergency admissions.

1.1.3 Differences between contemporary and retrospective PROMs

The current PROMs used in elective care asks patients to report their health status at different time points, and a change score is calculated from the difference between the pre-intervention (Q1 PROM) to the patient's post-intervention (Q2 PROM) score at three or six months, both based on contemporaneous ratings. This measure of change, based on the perspective held at the time of the assessment, assumes that the patients' perception of the construct under evaluation remains consistent between measurement points (the time interval between Q1 and Q2) [1].

In contrast, retrospective PROMs can be used when a contemporary Q1 score could not be anticipated or practically ascertained, i.e. before the emergency hospital admission. This is where the patient is asked to rate in retrospect how they thought they were at an earlier time.

Retrospective reporting from the patient's current perspective has been termed a 'then-test' by some researchers [29,30]. Respondents are asked to recall the point in time at which the contemporary pre-test Q1 PROMs would theoretically have been administered and to give a judgment of their level of functioning at the time. Thus, respondents are asked after the intervention or event, how they perceive themselves to have been beforehand. Retrospective self-reporting is extensively used in aetiological case-control studies and in cross-sectional surveys in which respondents are asked to recall characteristics of their health over a specified time frame which may be short (e.g. preceding week) or long (e.g. past year) [31].

These retrospective PROMs can be potentially influenced by recall bias [31]. A patient may not be able to recall their previous health, and may remember their health as being better or worse than what they would have reported at the time. Therefore recall bias has the potential to undermine the health change measured using retrospective reports.

Recall bias presents a threat to the reliability and credibility of studies using self-reported data. It arises when there are intentional or unintentional systematic differences in health as reported by patients at the time compared to how it is recalled at a later time point. The presence of any significant systematic differences in recall could lead to a misclassification of the health change measured (deterioration, no change, or improvement in the health status). If these systematic differences are not consistent between groups of patients, this may lead to underestimation or overestimation of the benefits gained from treatment [32,33].

1.2 Theoretical challenges

Several theoretical challenges complicate the use of retrospective PROMs for evaluating the impacts of emergency hospital care on patient outcomes. There are six main topics to consider as below.

1.2.1 The constructs to consider when seeking patients' reports of their health

When the use of PROMs questionnaires seeks patients' self-reports about their health, there is no precise definition or agreed terminology that defines what these instruments specifically measure. They are commonly referred to as measuring constructs such as quality of life, health-related quality of life, health status, functional status, and functional wellbeing.

Terms such as 'health status', 'health-related quality of life' and, sometimes more broadly, 'quality of life' are used interchangeably in the literature. As a result, some authors have noted that they lack real descriptive value [34]. Schipper et al suggested a simple definition to capture the meaning of health-related quality of life, 'the functional effect of an illness and its consequent therapy upon a patient, as perceived by the patient' [35]. The lack of agreed definitions to describe the constructs measured in PROMs also highlights the different focuses these instruments can have. These range from focusing predominantly on physical function e.g. mobility, to global questions on patients' health, and others that explore social and psychological factors.

The commonality of all instruments, noted in the Fitzpatrick et al review on the subject, is that they all address some aspect of the patient's subjective experience of health and the consequences of illness and treatment [36]. PROMs attempt to measure patients' views, feelings and experiences [37]. Patients are asked about their views regarding satisfaction, distress or symptom severity and these are unavoidably subjective in nature. The authors note and quote Albrecht that this subjective notion means that these questionnaires need not be objectively verified, even where questionnaire items ask for reports of very specific behaviours (for example, ability to walk a certain distance), and that they in principle can be objectively verified by separate clinical observation [36].

The inherently subjective nature of PROMs has previously been challenged in terms of robustness and scientific value as a form of evidence by some [38]. However, it is precisely this subjective focus of the patient that uniquely distinguishes PROMs from other forms of health outcome information.

When selecting PROMs for emergency patients, it is important to consider all the various constructs of health status (including physical, emotional/ psychological and cognitive domains) that would be important to capture. These domains may be interdependent in their effects on a patients' quality of life.

Lim et al have explored the relationships between the constructs of health status measures captured by SF-36 and EQ5D compared to domains that are seen as important to patients after critical care admission. Their study showed that PROMs questionnaires have different focuses on the constructs of health status and some are more complete in capturing certain constructs by the nature of their design [39].

The causes that lead to emergency admissions are multiple and complex and therefore the impacts on different areas that affect patient's quality of life are also likely to be multiple. Lim et al found in their study that concentration and memory were important cognitive constructs for critical care patients but capture of this by SF-36 may be incomplete [39]. This highlights that it is important when selecting PROMs to consider the dimensions of health status that would be most relevant to emergency admissions through involvement of patients and clinicians and then selecting validated tools that capture constructs seen as important to the respective patient groups. It is therefore important to achieve the right balance by including generic instruments that cover a breadth of constructs as well as condition specific instruments that deal with the clinically relevant domains in detail.

Psychometric theory involves assessing the measurement characteristics of scales and psychometric properties of validity, reliability and responsiveness, and this provides an important scientific basis for the selection of PROMs [40]. Development of clinically useful PROMs should take into account the following: i) Ensure that a specific purpose, focus, and setting are clearly identified. ii) Recognise that high statistical scores for reliability and validity may not be pertinent to a given situation or context. iii) Avoid indexes involving combinations of excessive numbers of variables. iv) Let patients choose the most significant foci and components of the indexes. v) Seek greater

communication and understanding among multidisciplinary collaborators, especially where there may be differences in the ethos and goals with which they approach the construction of health status [40]. Careful consideration of these qualities when selecting PROMs is equally important in both elective and emergency admissions, contemporary or retrospective.

1.2.2 The choice between two types of emergency admissions – sudden unexpected vs. exacerbation of long-term conditions

Emergency admissions are caused by a spectrum of conditions and illnesses. At one end, there is the sudden onset of an unexpected acute event that leads to a rapid deterioration of a patient's baseline health-related quality of life (e.g. road traffic accident), to the other end of the spectrum where the emergency admission is a result of an exacerbation of a long-term condition (e.g. chronic obstructive pulmonary disease) [41]. There are some that fall between the extremes in that the patient has a long-term condition, such as ischaemic heart disease, but the timing of an acute myocardial infarction is not predictable.

The method currently used in collecting elective surgery PROMs, (i.e. capturing a contemporaneous baseline prior to the intervention and then capturing a contemporaneous follow-up) is not possible in unexpected emergencies. Like elective surgery, the purpose of these hospital interventions for unexpected emergencies is to restore patients' health to their baseline. Similarly, the outcome of the patient's health is highly dependent on the quality of emergency hospital services and the care received during these inpatient episodes. Hence, to evaluate the impact of healthcare using PROMs, an alternative method to capture the baseline health status prior to the onset of the acute event is required for unexpected emergency admissions.

In contrast, contemporaneous baseline PROM capture is possible for patients that are admitted due to an exacerbation of a long-term condition. Studies have shown that PROMs can be captured as part of high quality patient centred care for people with long-term conditions [42,43]. When considering health-related quality of life over time, acute exacerbation in patients with long-term conditions' health-related quality of life mapped onto a graph may look more like undulating curves rather than sudden peaks and troughs. It is also

more ideal in these situations to monitor the effectiveness of the entire care pathway rather than the hospital episode for one particular emergency admission. Due to the range, diversity and intermittent nature of hospital admissions in patients with exacerbations of long-term conditions [42], it is difficult to attribute changes in PROMs in a meaningful way to care received during any particular admission. The hospital episode resulting from an exacerbation of a long-term condition may play a lesser role in the patient's overall health outcome compared with the input from primary care, outpatient and community services.

For this reason, the focus in this thesis is on whether PROMs can be used in sudden unexpected emergency admissions. This is also an area of increasing demand, resource use and currently where the NHS knows the least about the quality of care and outcomes other than mortality [44].

1.2.3 The appropriate point during the emergency admission to collect retrospective baseline PROM data

The role of PROMs in elective surgical procedures is relatively straightforward, with an aim to assess the effectiveness of discrete procedures in relation to patients with relatively clearly defined problems for which surgery is expected to be effective. Their role in evaluating emergency care is less straightforward, since patients typically have more complex and varied problems and interventions may be high risk.

The use of PROMs in unexpected emergencies provides an opportunity to gain information on patients' health-related quality of life, and compares the quality of health services in a similar manner to the use of PROMs in elective admissions. In contrast to the latter, the appropriate timing of PROMs collection in emergencies is particularly challenging to identify as the unexpected event that leads to the emergency admission can be complex to treat and cause multiple physical, social and emotional problems. The admission may involve several service providers and interventions, and recovery (whether full or partial) often takes many months.

Furthermore, the hallmark of high quality care for elective PROMs is usually seen as a positive health change from before and after PROMs questionnaires [9]. Health gains are more predictable and far more easily observed in elective interventions such as hip and knee replacements than for emergency care such as a hip fracture or heart attack. In contrast, in unexpected emergency admissions due to acute injury or sudden illness, the hallmark of high quality care and the result of health service inputs may not be a health gain but instead a restoration to pre-event level. In some cases, there may even be a health loss when compared to the patient's pre-event health-related quality of life, for example if the patient only achieves partial recovery.

There are many practical challenges facing collection of PROMs in emergency admissions at the time of the hospital admission, whether recalled or contemporary. Patients are acutely unwell and in pain making completing a questionnaire difficult. In addition, there may be insufficient time for patients to complete their questionnaires on admission before any urgent surgical or medical intervention. In fact, it would often be impractical to ask a patient to complete a PROMs questionnaire at the point of admission, when the event (acute injury or sudden illness) has reduced a persons' well-being to a state that warrants emergency medical treatment. In some instances, the patient may not even be conscious or cognitively well enough, and hence it would not be ethical or safe to divert clinical resources (e.g. staff time) on collecting PROMs when acute treatment is paramount. Therefore the point in time during the inpatient episode at which a retrospective PROM should be collected is not simply upon admission and must be considered and chosen carefully.

The time point which is identified as the patient's baseline health for unexpected emergencies may require a different approach to elective admissions. Consideration must be given to whether baseline should be measured immediately at the point of admission, or at a time before the acute event (e.g. one week before admission).

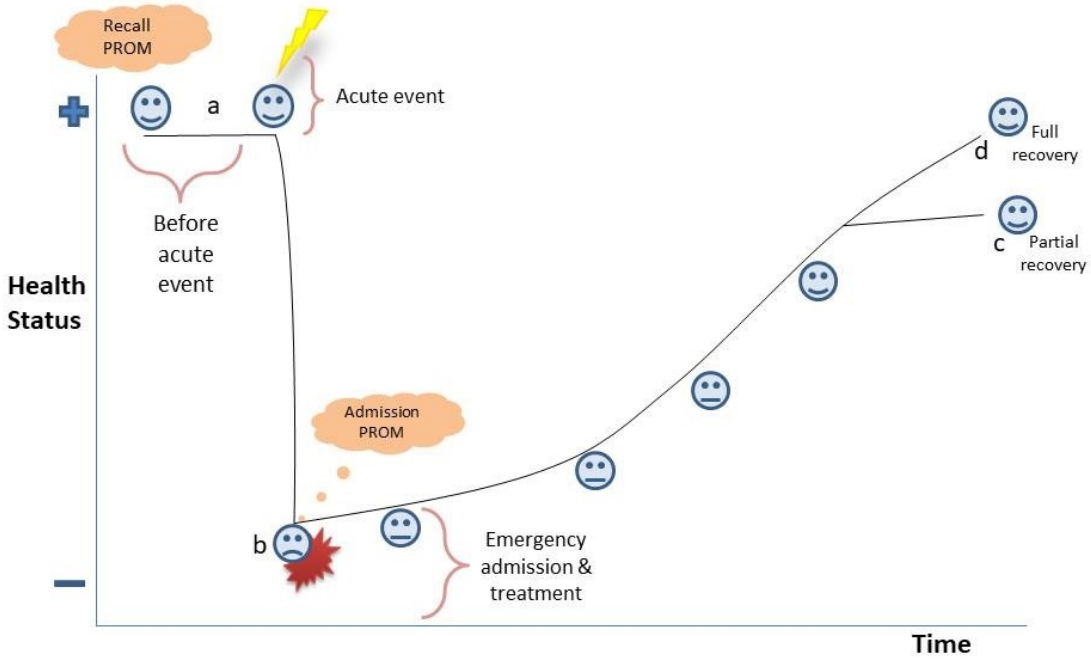
In theory, both the pre-event baseline health status (point *a* in *Figure 1*) as well as the acute emergency health status (*b*) would provide insight into health change for patients and could be used for comparison with a follow-up

questionnaire (point *c* or *d*). However, if the aim of capturing PROMs for unexpected emergency admissions is primarily to measure the effectiveness of health services, and the aim of healthcare is typically to restore a patient to their full potential with regards to their health-related quality of life, it is therefore a comparison with a patients' pre-event baseline (change between *c/d* and *a*) that should be used to determine the effectiveness of the health service.

Furthermore, a patient's health-related quality of life at the point immediately after an unexpected emergency event may be much worse than normal (change from *a* from *b*). Thus, judging an intervention or input by comparing *b* to *c* or *d* would demonstrate a larger difference in health change that is partly caused by the acute deterioration in health status caused by the patient's acute event/ condition, and could obscure the ability of using PROMs for comparisons in the effectiveness of healthcare provided. This risks negating any subtler differences in quality of care should emergency PROMs be used to support clinical quality benchmarking or comparisons between providers.

Measuring point *b* is problematic and any health status captured may mask the true assessment of the quality of health services. Therefore the recommended decision is to evaluate from point *a*, the pre-event baseline health status. This however is impossible to obtain prior to the acute event and therefore research into understanding the reliability of retrospective PROMs by patient recall to obtain a baseline health-related quality of life is particularly pertinent.

Figure 1-1 Trajectory of health-related quality of life in unexpected emergencies



1.2.4 The validity of using retrospective PROMs when they have been developed for contemporary use and the interpretation of differences between retrospective and contemporary PROMs

With retrospective PROMs, health status is established using patient recall. Recalling a prior health state that the patient is no longer experiencing can place considerable cognitive demand on the patient. One criticism is based on the contention that patients are unable to accurately recall prior health states, and recall is influenced by the patient's current symptoms [45]. There are three theories in cognitive psychology to consider regarding the use of recall in self-reported data, namely recall bias, response shift and implicit theories of change. These theories and their implication on both retrospective and contemporary PROMs data are discussed in turn.

Firstly, recall bias can be present in studies that use self-reported data, which are inherently subjective. It arises when there are systematic (intentional or unintentional) differences in health as reported by patients at the time compared to how it is recalled at a later time point [32]. If there are significant systematic differences in the way patients recall information compared to how it is reported at the time in retrospective PROMs, it would be important to investigate whether these were consistent and predictable across all individuals and groups, and the ways to correct or adjust for such when analysing and interpreting the data. If there is large differential recall (where the direction and extent of this recall bias randomly differs between individuals and patient groups) this could be more difficult to correct for and can introduce bias in comparisons of outcomes between different patient groups.

It is also important to recognise that often the boundaries between retrospective and contemporaneous PROMs are not straightforward. Contemporaneous PROMs can at times have a component of recall when questions are phrased to ask the patient to respond according to their perception of their health-related quality of life over a specified time period; e.g. in the Oxford Knee Score, the preceding four weeks. Often PROMs are developed and validated with the consideration of this recall period included in their guidance [46].

Secondly, some have argued that retrospective PROMs may be subjected to response shift, where a person's perception of their health changes over time as result of recalibration (result of a change in one's internal standards of measurement), reconceptualization (change in one's definition of the domains making up their health-related quality of life), and reprioritisation (a change in one's values on the importance of aspects making up their health-related quality of life) in their own self-appraisal. However, this phenomenon is not unique to retrospective PROMs and contemporary PROMs are also subject to this. In the literature, response shift theorists have at times supported the use of the retrospective ratings rather than contemporary ratings precisely for this reason, that the 'true' health change could be detected if the baseline was captured closer to follow-up by retrospective methods when the response shift has already occurred. For this argument to be justified any response shift present must be sustained; as otherwise, if the patient's experiences resulted in constant recalibration, then retrospective recall at one time point may not reflect that of another time point.

All PROMs are affected by theories that threaten internal validity. Authors have attempted to develop models to differentiate or isolate effects of response shift and recall bias. A model that has been used for detecting and quantifying response shift is a retrospective-minus-contemporary score developed by Schwartz and Sprangers called 'then-test minus pre-test', [47]. Patients are first asked to retrospectively re-evaluate their level at baseline from their current perspective ('then-test'). Recalibration response shift is represented by the 'then-test minus pre-test' difference score. The authors propose that if the contemporary (the authors call this 'post-test minus pre-test') difference score represents the reported treatment effect; the full treatment effect is represented by the 'post-test minus then-test' difference score due to the same internal standard of judgement [30].

When retrospective PROMs are used with a short (up to two weeks) recall period, the evidence suggests that the magnitude of any recall bias and response shift is small [48,49].

A further consideration is whether an event that changes a patient's health state suddenly (in unexpected emergency admissions) alters the way in which a prior health state is recalled. This leads onto the role of implicit change theory. This theory refers to the idea that the nature of recall is the retrieval of stored memory using a reconstruction of the past. The starting point of this reconstruction is based in the present, and may be prone to error, unless there is a significant event which the subject can use to anchor their basis of recollection. For example, a major surgery or catastrophic event may cause patients to anchor their memories, therefore providing a contextual reference point with which to associate their memory [50]. If a significant event occurs (e.g. major surgery), this may provide a sufficiently vivid event for people to anchor memories, and accurately recall their prior health status [48]; this could also be the case for patients with unexpected emergency admissions.

Therefore, all three theories can affect self-reported data including PROMs to certain extents, whether it is reported contemporaneously or retrospectively. Whilst it important to acknowledge that there are factors at play whenever information is asked of patients to recall or self-report, it is also important to credit that PROMs (whether retrospective or contemporary) are inherently subjective as they record patients' perception of their health. Therefore, it is the judgement made at the administered time point that ought to be considered the gold standard [36].

Since both contemporaneous PROMs and retrospective PROMs are subject to the influences of recall bias, response shift, and implicit theories of change, both types of PROMs should be intended for use in evaluation on the effectiveness of health services and clinical outcomes within their own respective contexts. Hence, the approach in the analysis presented in this thesis is to compare contemporary PROMs and retrospective PROMs, explore factors that could influence differences between them, and evaluate their degree of agreement as a measure of reliability. These factors are likely to be due to a combination of patient factors and other methodological factors, which may be specific to certain patient groups and some may be generalisable across patient groups whilst others may not. The factors identified could

contribute to any modelling required for adjusting data to make valid comparisons.

1.2.5 The generalisability of the relationship between a retrospective PROM and a contemporary PROM in elective surgery for emergency admissions.

The approach taken in the analysis presented in this thesis is to use evidence of agreement between contemporary PROMs and retrospective PROMs collected in elective patients to inform the use of the latter in patients receiving emergency care. This raises the question of whether evidence of differences between retrospective and contemporary PROMs in elective patients can be generalised to patients that experience unexpected injury or illnesses that lead to emergency admissions, Empirically, it is not possible to compare contemporaneous PROMs with retrospective PROMs directly in patient cohorts admitted with an emergency, but we can consider factors that may support or restrict generalisability between patient groups.

Runkel & McGrath suggested several types of generalisability for consideration [51]. One type is whether a specific treatment will produce the same results in different circumstances. Factors for consideration in the case of using PROMs in elective and emergency admissions are whether there are certain clinical factors and settings that could lead to divergence between elective and emergency care. The settings of elective surgery and unexpected emergencies are both in hospitals which limits this variability. It is however important to collect data in a number of different hospitals to investigate the range of results in different hospital environments.

A second and more pertinent type of generalisability concerns the subjects of the study. While the results of a study are internally valid for the subjects it tests, one should also consider whether there are other factors particular only to the characteristics of the study population and thus prevents the results from being generalised beyond that group [51]. Specifically, whether elective surgical patients differ as a group compared to emergency admission patients when it comes to the way they report and recall their health status. The characteristics of elective orthopaedic patients, when compared with acute

injured patients (for example), may differ in that elective patients may be older, and there may be a higher proportion of males in acute injury patients. It would therefore be important to explore the role patient characteristics play in recall.

There is very limited literature both on whether patients who are admitted unexpectedly report or recall their health status differently and whether this is significant compared with elective patients. However, there is evidence on the broader subject that we can draw insight from.

Studies have shown that with emergency admission patients, their recall of their pre-event baseline health remains stable over time, agreement between recalling pre-event health statuses at different time points remain high, and patients are able to recall similar pre-event health both during their hospital admission and post-discharge at home [37,52]. These studies suggest stability in recall ability in emergency admission patients. Studies in elective patients have also shown that recall is similarly stable over time [49,53]. Patients' characteristics such as age may contribute to differences in recall, and the extent of these effects is explored further in this thesis.

Another related aspect that offers some insight into this topic is the broader question of whether patients' valuations or health-related quality of life preferences are different from the general population. Although not directly compared between patients groups, these studies discuss relevant factors that might affect peoples' perception of their health-related quality of life in different circumstances.

Evidence suggests that patients assign valuations of their own actual health states differently when compared to general populations' hypothetical valuations [54]. A previous review by De Wit et al and a meta-analysis by Peeters et al have found that patients' valuations tend to be higher than hypothetical valuations of descriptions of the same states by the general population [55,56]. However, Dolders et al found these relationships were not significantly different overall and that age and gender did not contribute significantly to the difference between patient and population preferences [57]. Wilson et al studied acutely injured patients specifically and found that injured

patients valuations of their own health are slightly higher than the general population's hypothetical valuations of the same EQ-5D states [58].

Considering this broader perspective, although inconclusive, it appears that the literature weighs towards the suggestion that being a patient (being in that particular health state) can alter one's preferences and valuations. The question is whether these valuations also change between different patient groups, such as patients with a long-term condition, elective surgical patients, and acute injury patients.

Ubel et al discusses some of the potential factors leading to discrepancies in valuation between patients and the public; possibilities include: i) the two groups are valuing different health states although they are represented identically on a health utility measure (e.g. when aspects of health are not captured fully by the questionnaire or description of the health state). ii) Patients adapt to their poor health states (and so value them relatively higher). iii) The 'focusing illusion', whereby members of the general population over-emphasise aspects of health-related quality of life most affected by illness. iv) A shift of reference points as people assess health states with reference to their current health [59].

At least three, (i) (ii) and (iv) of the four points above could potentially be a factor for differences between elective and emergencies patient groups. For example, patients who have suddenly developed an unexpected emergency in a short space of time may be more similar to the general public in their perspective of health preferences compared to elective surgical patients who may have had their condition for some time prior to their operation and therefore may have adapted to their poorer health state. This is however difficult to conclusively demonstrate. There is at least the potential for unexpected emergency admission patients to report their health-related quality of life and health status differently from elective patients.

One cannot conclusively know whether the way patients' valuation of retrospective recall in health statuses would differ between elective and emergency patients. This would, however, not affect the use of PROMs within

the emergency context in monitoring effectiveness or for comparing providers. It could pose a potential issue if PROMs were used in a way for resource allocation in economic evaluations comparing between elective and emergency conditions.

Both acceptability and feasibility are also two other important aspects of generalisability for PROMs [36] that can differ from elective admission contexts. It is important to explore these criteria in greater depth in unexpected emergencies.

Firstly, acceptability, the extent to which an instrument is acceptable to patients. Indicators of acceptability include administration time, recruitment and response rates. Factors such as the mode of administration, questionnaire design, and the health status of respondents that affect acceptability should be explored [36]. Issues of acceptability were considered in discussions with relevant stakeholders in selected emergency conditions, selecting questionnaires that are acceptable for use by the intended data collection method.

Secondly, exploring the feasibility criteria; testing the ease of administration and processing of an instrument for the specific context is particularly important for unexpected emergency admissions. Instruments should be selected based on ease of administration and minimal disruption to clinical care. Other factors to explore with regards feasibility include recruitment differences according to condition and sites, and response biases by patient characteristics and the interpretability of the findings.

1.2.6 Generalisability of two sudden emergency events to other emergency causes

For pragmatic reasons, it would not be practical to test the feasibility of PROMs collection in all unexpected emergency admissions. I have therefore identified two sentinel conditions within unexpected emergencies in this thesis.

Generalisability to other emergency causes is important to consider, based on Runkel & McGrath's aspects of generalisability for consideration (as discussed earlier in topic 5) [51]. It is important to find out if it is equally feasible to collect

PROMs, and whether the health change measure by PROMs is consistent for a range of different types of unexpected emergency admissions. It is therefore paramount to incorporate emergency conditions from both surgical and medical areas, leading to my choice of emergency laparotomy and acute myocardial infarction. It is also important to conduct a feasibility study in a variety of hospital settings, involving different staff, patients, and environments.

1.3 Aim and objectives

1.3.1 Aim

To identify a reliable method for routinely collecting PROMs in emergency admissions to hospital.

1.3.2 Research Objectives

- To review the literature on the relationship of PROMs acquired retrospectively with PROMs acquired contemporaneously and from population surveys
- To compare retrospective and contemporary PROMs in two elective conditions
- To compare retrospective PROMs and population values in two elective conditions
- To test the feasibility of collecting retrospective PROMs in two emergency admissions
- To make recommendations on a reliable method for determining baseline PROMs in emergency hospital admissions

1.4 Overview of the structure to the thesis

There are 8 chapters in my thesis that describes the studies that have been conducted to address the aims and objectives. I have answered each separate objective in the following chapters, alongside original manuscripts that I have prepared for publication at each stage.

Chapter 2 (Paper1) presents a narrative synthesis review of what is known about the use of retrospective PROMs and their reliability in terms of agreement when compared to contemporary PROMs and compared its use to PROMs from population surveys. It discusses findings about the factors that influence recall of prior health status, and situations where agreement is strong. This chapter informs the methodologies used for my subsequent study designs.

Chapter 3 (Paper 2) describes a longitudinal cohort study conducted in four NHS hospitals with elective patients to directly compare contemporary PROMs and retrospective PROMs. This work was conducted in patients undergoing elective hip and knee arthroplasty surgery in whom contemporary PROMs are already collected. Patients who were participating in the NHS National PROMs Programme were recruited to the study during their inpatient period following their elective surgery. Consented patients then completed a retrospective PROM, where they were asked postoperatively about how they were before their hospital admission. Their retrospective and contemporary PROMs questionnaires were linked to compare the agreement between the scores.

Chapter 4 describes a comparison of retrospective and contemporary PROMs and population values from the General Practice Patient Survey (GPPS) in partnership with the University of Exeter. For this chapter, I explored the use of different matching techniques (using different patient characteristics criteria) for surgical patients from the elective study cohort (Chapter 3) to the GPPS population. These exploratory methods were conducted in order to investigate whether the use of already collected EQ-5D population values could be harnessed for use as a surrogate of patients' baseline in place of a retrospective PROM and discussed how these methods could be applied and investigated in the future with emergency patients cohorts.

Chapters 5 - 7 (Papers 3 - 5) describe a feasibility study of using retrospective PROMs methodology to collect baseline PROMs, as well as a follow-up PROMs to assess the short-term outcomes, in two contrasting reasons for emergency admissions: in patients undergoing emergency laparotomy (EL) for gastrointestinal conditions (excluding appendicitis) and those with ST-elevation myocardial infarction (STEMI) who undergo an emergency percutaneous coronary intervention (PCI). This study explored the acceptability and interpretability aspects of feasibility in the context of using PROMs during emergency admissions. It also explored the use of retrospective PROMs to collect baseline data in emergency admissions in terms of recruitment and response rates achieved in a variety of different hospitals, and in two contrasting patient/ disease groups.

My final chapter, Chapter 8 provides an overview of the main findings and discusses the limitations of the thesis, as well as the need for future research, along with implications how this research would inform practice and policy.

1.5 Contribution of the candidate to the thesis

I undertook the narrative literature review and took the lead in the planning of the study design, securing NHS ethical approval and NHS digital data applications of all studies which make up this thesis and was supported in this by my supervisor Professor Nick Black and co-authors. I planned, and trained all local site leads and their teams for data collection for the three cohort studies in 20 different trusts and acted as the Chief Investigator for each of these, liaising regularly during the data collection phases to ensure the smooth running of the study. I led and conducted the feasibility studies' follow-up PROMs questionnaire data collection from LSHTM by mail with the support of a part-time administrative assistant Mrs Christina Breach whom I trained. I undertook all the data analysis and was provided with statistical support by my second supervisor, Dr Jenny Neuburger. Professor Nick Black provided guidance on presentation of the findings of the research papers in this thesis. I produced the first draft of each research paper and made changes in response to co-authors' feedback.

My PhD was funded by the Economic and Social Research Council and affiliated to the National Institute of Health Research (NIHR), Collaborations for Leadership in Applied Health Research and Care (CLAHRC) North Thames. The candidate was the Chief Investigator; Professor Nick Black was PhD supervisor in all NHS Integrated Research Approval System (IRAS) applications for the studies. Dr Dave Murray and Professor Steffen Petersen acted as clinical collaborators for the feasibility studies.

1.6 References

1 Department of Health. Patient Reported Outcome Measures (PROMs) in England A Methodology for Identifying Potential Outliers. 2009.

2 Black N. Patient reported outcome measures could help transform healthcare. *BMJ* 2013;**346**:19–21.

3 Greenhalgh J, Long AF, Flynn R. The use of patient reported outcome measures in routine clinical practice: lack of impact or lack of theory? *Soc Sci Med* 2005;**60**:833–43.

4 Morris S, Devlin N PD. *Economic Analysis in Health Care*. Chichester: : John Wiley & Sons 2007.

5 Department of Health. NHS Outcomes Framework: at-a-glance List of outcomes and indicators in the NHS Outcomes Framework for 2016-17. 2016.

6 Department of Health. The Government ’s revised mandate to NHS England for 2017-18. 2018.

7 Browne J, Jamieson L, Lawsey, J, van der Meulen J, Black N, Cairns J, Lamping D, Smith, S, Copley L, Horrockes J (2007). Patient Reported Outcome Measures (PROMs) in Elective Surgery. 2007.

8 Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;**35**:1095–108.

9 Devlin NJ, Appleby J. The King's Fund. Getting the most out of PROMs: Putting health outcomes at the heart of NHS decision-making. 2010. http://www.kingsfund.org.uk/publications/proms.html

10 Szende A, Oppe M DN. *EQ-5D Valuation Sets: An Inventory, Comparative Review and Users’ Guide*. Springer Netherlands 2007.

11 Dawson J, Fitzpatrick R, Murray D, *et al.* Comparison of measures to assess outcomes in total hip replacement surgery. *Qual Health Care* 1996;**5**:81–8.

12 Greene ME, Rolfson O, Gordon M, *et al.* Standard Comorbidity Measures Do Not Predict Patient-reported Outcomes 1 Year After Total Hip Arthroplasty. *Clin Orthop Relat Res* 2015;**473**:3370–9.

13 The NHS Confederation. System under strain. 2018.

14 Taunt R, Lockwood A, Berry N. More Than Money: Closing The NHS Quality Gap. 2014.

15 Institute of Medicine (US) Committee on Quality of Health Care in America. *Crossing the Quality Chasm: A New Health System for the 21st Century*. 2001.

16 Black N. Quality improvement in the NHS. *BMJ* 2008;**336**:1143 LP-1143.

17 Department of Health. High Quality Care For All NHS Next Stage Review Final Report. 2008.

18 Department of Health. NHS Outcomes Framework 2013 to 2014. 2012. https://www.gov.uk/government/publications/nhs-outcomes-framework-2013-to-2014

19 Greenhalgh J, Dalkin S, Gooding K, *et al.* Functionality and feedback: a realist synthesis of the collation, interpretation and utilisation of patient-reported outcome measures data to improve patient care. *Heal Serv Deliv Res* 2017;**5**.

20 Neuburger J, Hutchings A, Allwood D, *et al.* Sociodemographic differences in the severity and duration of disease in patients undergoing hip or knee replacement surgery. *J Public Heal* 2012;**34**:421–9.

21 Barnsley P, Towse A, Schaffer SK, *et al.* Office of Health Economics. Critique of CHE Research Paper 81: Methods for the Estimation of the of the NICE Cost Effectiveness Threshold. 2013.

22 Claxton K, Martin S, Soares M, *et al.* Centre for Health Economics University of York. Methods for the Estimation of the NICE Cost Effectiveness Threshold. 2013.

23 Browne JP, Cano SJ, Smith S. Using Patient-reported Outcome Measures to Improve Health Care. *Med Care* 2017;**55**:901–4.

24 Steventon A, Deeny S, Friebel R, *et al.* The Health Foundation. Emergency hospital admissions in England: which may be avoidable and how? 2018.

25 National Audit Office. Reducing Emergency Admissions Report to Dept of Health & Social Care NHS England. 2011.

26 Tian Y, Dixon A, Gao H. The King's Fund. Data briefing - Emergency hospital admissions for ambulatory care-sensitive conditions: identifying the potential for reductions. 2012.

27 Jenkinson C, McGee H. *Assessment and evaluation of health and medical care: a methods text*. Open University Press 1997.

28 NHS England. GPPS Surveys and Reports.2017. www.gpatient.co.uk/download/Questionnaires/

29 Visser M, Oort F, Sprangers M. Methods to detect response shift in quality of life data: a convergent validity study. *Qual Life Res* 2005;**14**:629–639.

30 Sprangers MA, Van Dam FS, Broersen J, Lodder L, Wever L, Visser MR, Oosterveld P SE (1999. Revealing response shift in longitudinal research on fatigue--the use of the thentest approach. *Acta Oncol (Madr)* 1999;**38**:709–18.

31 Hennekens CH, Buring JE. *Epidemiology in Medicine*. Lippincott Williams & Wilkins 1987.

32 Grimes DA, Schulz KF. Bias and causal associations in observational research. *Lancet (London, England)* 2002;**359**:248–52.

33 Basso O, Olsen J, Bisanti L, *et al.* The performance of several indicators in detecting recall bias. European Study Group on Infertility and Subfecundity. *Epidemiology* 1997;**8**:269–74.

34 Spitzer WO. State of science 1986: Quality of life and functional status as target variables for research. *J Chronic Dis* 1987;**40**:465–71.

35 Bullinger M, Quitmann J. Quality of life as patient-reported outcomes: principles of assessment. *Dialogues Clin Neurosci* 2014;**16**:137–45. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4140508/

36 Fitzpatrick R, Davey C, Buxton MJ, *et al.* Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 1998;**2**:i–iv, 1-74.

37 Mor V, Guadagnoli E. Quality of life measurement: a psychometric tower of Babel. *J Clin Epidemiol* 1988;**41**:1055–8.

38 Deyo RA. THe quality of life, research, and care. *Ann Intern Med* 1991;**114**:695–7.

39 Lim WC, Black N, Lamping D, *et al.* Conceptualizing and measuring health-related quality of life in critical care. *J Crit Care* 2016;**31**:183–93.

40 Hays RD, Anderson R, Revicki D. Psychometric considerations in evaluating health-related quality of life measures. *Qual Life Res* 1993;**2**:441–9.

41 National Audit Office. Emergency admissions to hospital: managing the demand. *Hc* 2013;**739**:4.

42 Peters M, Crocker H, Dummett S, *et al.* Change in health status in long-term conditions over a one year period: a cohort survey using patient-reported outcome measures. *Health Qual Life Outcomes* 2014;**12**:123.

43 Stiggelbout AM, Van der Weijden T, De Wit MPT, *et al.* Shared decision making: really putting patients at the centre of healthcare. *BMJ* 2012;**344**:e256.

44 Purdy S. Avoiding hospital admissions. London The Kings Fund: 2010.

45 Herrmann D. Reporting current, past, and changed health status. What we know about distortion. *Med Care* 1995;**33**:AS89-94.

46 U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes* 2006;**4**:79.

47 Sprangers M. Response-shift bias: a challenge to the assessment of patients' quality of life in cancer clinical trials. *Cancer Treat Rev* 1996;**22 Suppl A**:55–62.

48 Bryant D, Norman G, Stratford P, *et al.* Patients undergoing knee surgery provided accurate ratings of preoperative quality of life and function 2 weeks after surgery. *J Clin Epidemiol* 2006;**59**:984–93.

49 Howell J, Xu M, Duncan CPCP, *et al.* A Comparison Between Patient Recall and Concurrent Measurement of Preoperative Quality of Life Outcome in Total Hip Arthroplasty. *J Arthroplasty* 2008;**23**:843–9.

50 Ross M. Relation of implicit theories to the construction of personal histories. Psychol Rev 1989;96:341e57.

51 Runkel PJ, McGrath JE. *Research on Human Behavior: A Systematic Guide to Method.* American Educational Research Association 1973.

52 Williamson ODOD, Gabbe BJBJ, Sutherland AMAM, *et al.* Does recall of preinjury disability change over time? *Inj Prev* 2013;**19**:238–43.

53 Marsh J, Bryant D, MacDonald SJ. Older patients can accurately recall their preoperative health status six weeks following total hip arthroplasty. *J Bone Jt* 2009;**91**:2827–37.

54 Kostopoulou O. The transient nature of utilities and health preferences. *Med Decis Mak* 2006;**26**:310–312.

55 Peeters Y, Stiggelbout AM. Health state valuations of patients and the general public analytically compared: a meta-analytical comparison of patient and population health state utilities. *Value Heal J Int Soc Pharmacoeconomics Outcomes Res* 2010;**13**:306–9.

56 De Wit GA, Busschbach JJ, De Charro FT. Sensitivity and perspective in

the valuation of health status: whose values count? *Health Econ* 2000;**9**:109–26.

57 Dolders MGT, Zeegers MPA, Groot W, *et al.* A meta-analysis demonstrates no significant differences between patient and population preferences. *J Clin Epidemiol* 2006;**59**:653–64.

58 Wilson R, Hansen P, Langley J, *et al.* A comparison of injured patient and general population valuations of EQ-5D health states for New Zealand. *Health Qual Life Outcomes* 2014;**12**:21.

59 Ubel PA, Loewenstein G, Jepson C. Whose quality of life? A commentary exploring discrepancies between health state evaluations of patients and the general public. *Qual Life Res* 2003;**12**:599–607.

Chapter 2

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Esther Kwong
Principal Supervisor	Nick Black
Thesis Title	The use of patient reported outcome measures (PROMs) for evaluating emergency admissions

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?	Journal of Clinical Epidemiology.		
When was the work published?	2017		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	Please refer to details on the following page

Student Signature: [redacted] **Date:** 24 August 2018

Supervisor Signature: [redacted] **Date:** 24 August 2018

2.1 Chapter 2

Retrospectively patient-reported pre-event health status shows strong association and agreement with contemporaneous reports

In this chapter, I report on a review of the literature to understand what is already known about the use of retrospective PROMs when compared with contemporary PROMs, and conducted a narrative synthesis on the knowledge on the agreement between these. I also reviewed studies that have compared retrospective PROMs with PROMs collected in population surveys. I conducted the literature review design, methods, and analysis independently with supervision from Professor Nick Black. The findings and results have been prepared as a first draft of the manuscript, with comments on drafts from Professor Nick Black. This was published in the Journal of Clinical Epidemiology.

ELSEVIER

Journal of Clinical Epidemiology 81 (2017) 22–32

Journal of Clinical Epidemiology

Retrospectively patient-reported pre-event health status showed strong association and agreement with contemporaneous reports

Esther Kwong*, Nick Black

Department of Health Services Research & Policy, London School of Hygiene & Tropical Medicine, 15-17 Tavistock Place, London WS1H 9SH, UK

Accepted 5 September 2016; Published online 10 September 2016

Abstract

Objective: The unpredictability of the occurrence of illnesses and injuries leading to most emergency admissions to hospital makes it impossible prospectively to collect preadmission patient-reported outcome measures (PROMs). Our aims were to review the evidence for using retrospective PROMs to determine pre-event health status and the validity of using general population norms instead of retrospective PROMs.

Study Design and Setting: Searches of Medline, PsycINFO, Embase, Global Health, and Health Management information. Six studies met the inclusion criteria for the first aim, and 11 studies addressed the second aim. Narrative syntheses were conducted.

Results: Strong associations were found between retrospective and contemporary PROMs in 21 of 30 comparisons (correlation coefficients over 0.68) and 20 of 24 showed strong agreement for continuous measures (intraclass correlations over 0.75). Categorical measures revealed only fair to moderate levels of agreement (kappa 0.3–0.6). Associations were stronger for indices than for individual items and for shorter time intervals. The direction of differences was inconsistent. Retrospective PROMs reported by elderly patients were similar to the general population but younger adults had been healthier.

Conclusion: Retrospective collection offers a means of assessing PROMs in unexpected emergency admissions. However, further research is needed to establish the best policy for their use.

Keywords: Patient-reported outcome measures; Health-related quality of life; Retrospective; Population norms; Recall bias; Response shift

1. Introduction

The growing acceptance of the importance of patients' views of their outcome when evaluating interventions and assessing the quality of services means that it is necessary to devise ways in which accurate patient-reported outcome measures (PROMs) can be obtained (referred to as PROs in the United States) [1]. PROMs are self-completed questionnaires where patients are asked to report their own state of health (multidimensional symptoms, functional status) and health-related quality of life (HRQL) at one point in time. PROMs can be categorized as generic (e.g., EuroQuol-5D, Short Form [SF]-36) or disease specific (Oxford Hip Score or Western Ontario and McMaster Osteoarthritis Index [WOMAC]). Generic PROMs capture broad domains on function or HRQL, can be converted into utility scores, and provide the means to compare between conditions and treatments. Disease-specific PROMs have greater sensitivity by incorporating aspects of function and HRQL specific to that condition [1]. By comparing measurements before and after a healthcare intervention, the outcome of care can be determined.

Emergency admissions make up 34% of hospital admissions in England [2]. They can be categorized as either a largely unexpected acute event, such as an acute myocardial infarction, stroke, or injury (about 70% of all emergency admissions) or as an exacerbation of an existing long-term conditions as occur in conditions such as diabetes or chronic obstructive pulmonary disease. Although these are not a clear-cut dichotomy, the two categories present different challenges when using PROMs. Unlike elective admissions when a PROM can be collected before treatment to capture the baseline health at the time (a contemporary PROM), for unexpected emergency admissions, this is not possible (This need not be a problem for emergency admissions due to exacerbations of long-term

Funding: E.K. is funded by an Economic & Social Research Council doctoral fellowship, grant reference ES/J500021/1.

* Corresponding author. Tel.: +44(0)20 7958 8285.

E-mail address: esther.kwong@lshtm.ac.uk (E. Kwong).

http://dx.doi.org/10.1016/j.jclinepi.2016.09.002

What is new?

Key findings

- There is a strong association between patient-reported outcome measures (PROMs) collected retrospectively and contemporaneous collection among patients undergoing elective surgery.
- Agreement is also strong for PROMs that are continuous measures but only fair to moderate for categorical measures.
- Retrospectively collected data suggest that young adult trauma patients are healthier than population norms. The reverse may be true for older patients admitted for medical conditions.

What this adds to what was known?

- Retrospective collection offers a means of assessing patient-reported outcomes in unexpected emergency admissions.

What is the implication and what should change now?

- Further methodological research is needed to establish the best policy for their use.

conditions, such as chronic obstructive pulmonary disease, when PROMs could be collected as part of their routine clinical management, i.e., a contemporary PROM). Therefore, for unexpected admissions, other methods must be used to assess patients' preadmission baseline health status.

There are two possible approaches. First, there is the use of retrospective PROMs, in which patients are asked to recollect (after their unexpected emergency event, such as an acute myocardial infarction) what their health status and quality of life was like just before the emergency event. This takes the place of contemporaneous collection before the event that can be done when considering planned elective treatments such as hip replacements. Retrospective self-reporting has been extensively used in etiological case–control studies and in cross-sectional surveys [3] in which respondents are asked to recall characteristics of their health over a specified time frame which may be short (e.g., preceding week) or long (e.g., past year).

Second, and much cheaper than retrospective reporting, is to use age–sex standardized PROMs which have been collected from the general population (or an appropriate comparison group) as part of a cross-sectional survey, as a surrogate measure of a patients' pre-event baseline health [4]. The use of population norms assumes that patients experiencing an emergency admission are typical of the wider population. This assumption could lead to an overestimate or underestimate of the impact of a health care intervention. If patients are in fact healthier at baseline than the general population (as might be the case when studying recovery from trauma that occurred while undertaking a fitness based sport such as rock-climbing), using the population norm as a surrogate baseline could lead to an "overestimate" of the treatment effect. On the other hand if patients were in worse health than their peers beforehand (as might be expected for those suffering a heart attack), an "underestimation" of the treatment effect will be observed.

Although there has been no review of the strength of association and of agreement between these two approaches in emergency admissions, two systematic reviews have considered other aspects of recall. One considered the length of recall periods for PROMs in clinical trials and concluded that the optimum depended on two broad categories of factors: characteristics of the phenomenon being recalled (such as how recently it had occurred, its attributes, its complexity) and the context of the recalled phenomenon (such as its salience, the patient's mood) together with the nature of the topic [5]. The second review concluded that recall bias is a concern with PROMs and called for more research to understand and identify situations where the use of recall is acceptable [6].

Our aims were to review systematically the scientific evidence on (1) the extent of association and agreement between PROMs collected retrospectively and contemporaneously to determine pre-event health status and HRQL and (2) the validity of using general population norms for determining the pre-event health status and HRQL of people with an unexpected emergency admission to hospital.

2. Study design and settings

2.1. Literature search

A search was conducted on studies either (1) comparing retrospective and contemporary PROMs (health status, symptoms, functional status, HRQL) or (2) comparing retrospective PROMs and population norms. For inclusion, studies had to be in English; involve self-completed questionnaires; have a recall period of no more than 6 months. In addition, for comparisons of retrospective and contemporary PROMs, studies had to include a quantitative estimation of the strength of association (Pearson or Spearman rank correlation) or agreement (intraclass correlation coefficient or kappa score). No additional analyses were undertaken to determine missing correlations or levels of agreement.

Our focus was on methods for estimating patients' pre-event health or HRQL that could be used to determine the extent to which treatment restored them to their previous state of health. Many studies ask patients themselves to assess the extent of change in their health (single transitional items) [7,8], but this is a different methodological

approach to that of comparing assessments at two points in time and were therefore excluded from this review.

Five databases were searched: Medline, PsycINFO, Embase, Global Health, and Health Management information. A free-text search strategy was used as subject headings were too broad and nonspecific for the research question. The detailed concepts, keywords, and search terms are summarized in Table 1, and the complete search strategy is summarized in Table 2. A forward and backward snowballing strategy was used to complement the free-text search.

Identified articles were exported to a reference manager (Mendeley Desktop version 1.13) and duplicates removed. The title and abstracts were screened by one author (E.K.) to assess suitability. Studies in children, adolescents, carer proxies, and those with cognitive impairments were excluded. The remaining articles were read, and forwards and backwards searching of references was conducted (Fig. 1).

2.2. Quality appraisal

For studies comparing retrospective and contemporary PROMs, their methodological quality was appraised by one author (E.K.) using five relevant items selected from the Quality Appraisal of Diagnostic Reliability (QAREL) checklist [9]. These items cover the representativeness of participants, time interval between assessments, correct application of assessment, and appropriate statistical analysis. The other items were not applicable in this review: whether participants were blinded to their initial assessment, to other participants' assessments, to any reference standard, or to clinical information, or blinded to additional cues that were not part of the test. A simple summation of the five included items was calculated (0 = weak, 5 = strong). Given the heterogeneity of the studies in this review, a narrative synthesis was carried out.

2.3. Definition of strength of association and agreement

Association according to Pearson or Spearman correlation coefficients was classified as weak (below 0.36), moderate (0.36−0.67), strong (0.68−0.90), and very strong (above 0.90) [10].

Agreement according to intraclass correlation coefficients was classified as weak (below 0.36), moderate (0.36−0.67), strong (0.68−0.90), and very strong (above 0.90). Agreement according to kappa scores were classified as: slight (<0.20), fair (0.20−0.40), moderate (0.41−0.60), substantial (0.61−0.80), and almost perfect (0.81−1.0) [11].

3. Results

3.1. Search findings

Two hundred seventy-five articles were identified on Medline, 350 on Embase, 102 on PsycINFO, 18 on Global Health, and 2 on Global Management Information (all accessed 22 April, 2015). Having removed duplicates, 450 abstracts were reviewed of which four comparing retrospective and contemporary PROMs, and five comparing retrospective PROMs and population norms, met the inclusion criteria. Most of the studies were excluded either because they did not capture a contemporary baseline PROM measurement or there was no statistical assessment of the strength of association or agreement between contemporary and retrospective PROMs. A citation search on PubMed (forward and backward snowballing) identified 2 additional studies comparing retrospective and contemporary PROMs and six comparing retrospective PROMs and population norms (Fig. 1). All studies comparing retrospective and contemporary PROMs were methodologically strong according to the QAREL checklist.

3.2. Comparison of retrospective with contemporary PROMs

Of the six studies, one was from the United Kingdom [12], one was multinational [13], three were from Canada [14−16], and one from the United States [17] (Table 3). The studies involved 75−177 patients, with one exception with 770 patients [13]. Four involved patients with hip and knee problems and two were based on urological patients. Several reported on the level of agreement between retrospective and contemporary reports for more than one PROM.

Table 1. Literature search: concepts, keywords, and search terms

	Search terms			
Concepts	Retrospective	Population norms	Patient reported	Outcomes
Keywords	Retrospective	Population norm$	Self-report$	Outcome$
	Recall		Patient report$	Quality * life H?Q?L
	Historical		Patient recall$	EQ-5D function$
	Bias		Self-recall$	SF-36
	Recollected			Health status
				Symptom$

Abbreviations: EQ-5D, EuroQuol-5D; SF-36, Short Form−36.

Table 2. Search strategy

1. Retrospective or recall or historical or recollected
2. Bias
3. Population norm$
4. Self-report$ or patient report$ or patient recall$ or self-recall$
5. Outcome$ or quality * life or H?Q?L or EQ-5D or function$ or SF-36 or health status or symptom$
6. 1 OR 2 OR 3
7. 6 ADJ5 4
8. 7 ADJ10 5
9. Limit 8 to (humans)
Combined search string: ((retrospective or recall or historical or bias or population norms or recollected) adj5 ((self-report$ or patient report$ or patient recall$ or self-recall$) adj10 (outcome$ or quality * life or H?Q?L or EQ-5D or function$ or SF-36 or health status or symptom$))).mp.

Abbreviations: EQ-5D, EuroQuol-5D; SF-36, Short Form–36.

Eleven different PROMs were used including the SF-36 or SF-12 (four studies), the WOMAC (three studies), the American Urological Association Symptom Index (two studies), the Western Ontario Meniscal Evaluation Tool, the Knee Injury and Osteoarthritis Outcome Score, Oxford Hip Score, Lower Extremity Functional Scale, and the feeling thermometer. The time period for retrospective reporting was predominantly 2 weeks to 3 months although one study reported 3 days (in addition to longer periods) and one used 6 months.

All six studies assessed the level of association between retrospective and contemporary PROMs scores using correlation coefficients (four used Pearson and two used Spearman coefficients), all reported on the level of agreement (three used kappa statistics and three used intraclass coefficients). Most presented analyses of the full index scores though some reported on subscales. A total of 30 correlations coefficients of full-scale or subscale scores were reported, of which nine were moderate, 18 were strong, and three were very strong.

Three studies that each used several PROMs at different time points thus generating 24 comparisons, the level of agreement for continuous data (intraclass correlations) was very strong for eight, strong for 12 and four were moderate [14–16]. In contrast, for PROMs that were converted to categorical variables for analysis, kappa statistics revealed only fair to moderate levels of agreement [12,13,17].

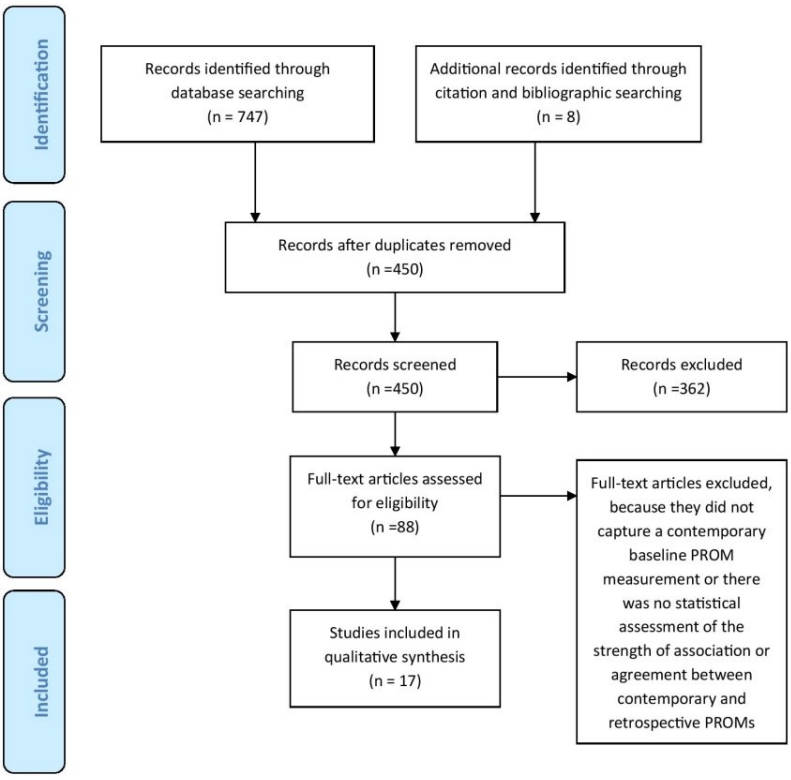


Fig. 1. Search results. PRISMA 2009 flow diagram. PROMs, patient-reported outcome measures. From Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 2009;6(7):e1000097; http://dx.doi.org/10.1371/journal.pmed1000097.

Table 3. Studies comparing retrospective to contemporary PROMs

Author Country/Year	Condition/procedure Recall period Sample size	PROMs	Level of association (correlation coefficient)	Level of agreement	Retrospective health compared to contemporary report[a]
Emberton [12] UK 1995	Benign prostatic hyperplasia 3 mo $n = 75$	AUA Symptom Index AUA Symptom Impact Index	*Pearson* Symptom Index: 0.6 Symptom Impact Index: 0.6	*Weighted kappa* Symptom Index: 0.3 Symptom Impact Index: 0.3	No difference
Lingard [13] USA, UK, and Australia 2001	Total knee arthroplasty 3 mo $n = 770$	Western Ontario & McMaster Osteoarthritis Index (WOMAC) pain scale SF-36 function scale	*Spearman* WOMAC (pain scale): 0.53 SF-36 (function scale): 0.48	*Weighted kappa* Individual items: 0.20–0.41	Worse for WOMAC pain scale (51.9% no difference, 31.3% recalled more pain, 16.8% recalled less pain) ($P < 0.001$) No consistent difference for SF-36 function scale (75% no difference, 11.8% recalled less limitation, 3.5% recalled more limitation) Patients recalled significantly less limitation for walking >1 mile ($P < 0.001$) but significantly more limitation for walking 100 yards ($P = 0.009$).
Bryant [14] Canada 2006	Knee surgery 2 wk $n = 177$	SF-36 International Knee Documentation Committee (IKDC) Subjective Form Anterior Cruciate Ligament–Quality of Life (ACL-QOL) Western Ontario Meniscal Evaluation Tool (WOMET) Knee Injury and Osteoarthritis Outcome Score (KOOS)	*Pearson* SF-36 (PCS): 0.81 SF-36 (MCS): 0.68 IKDC: 0.92 ACL-QOL: 0.86 WOMET: 0.88 KOOS: 0.93	*Intraclass coefficient* SF-36 (PCS): 0.81 SF-36 (MCS): 0.67 IKDC: 0.92 ACL-QOL: 0.86 WOMET: 0.88 KOOS: 0.93	No difference
Howell [15] Canada 2008	Total hip arthroplasty 3 days; 6 and 12 wk $n = 104$	WOMAC OHS SF-12 (PCS) SF-12 (MCS)	*Spearman* 3 days; 6 wk; 12 wk WOMAC: 0.80, 0.78, 0.86 OHS: 0.82, 0.80, 0.92 SF-12 (PCS): 0.66, 0.54, 0.76 SF-12 (MCS): 0.77, 0.71, 0.76	*Intraclass correlation* 3 days; 6 wk; 12 wk WOMAC: 0.86, 0.88, 0.93 OHS: 0.91, 0.88, 0.96 SF-12 (PCS): 0.83, 0.77, 0.90 SF-12 (MCS): 0.86, 0.84, 0.93	3 days: worse (OHS$\Delta = 1.58$ $P = 0.01$, WOMAC $\Delta = -2.21$ $P = 0.029$, SF-12 MCS $\Delta = -4.82$ $P < 0.001$) 6 wk: worse (SF-12 MCS$\Delta = -2.79$ $P = 0.01$) 12 wk: no difference
Marsh [16] Canada 2009	Total hip arthroplasty 6 wk $n = 174$	WOMAC OHS SF-12 (PCS) SF-12 (MCS)	*Pearson* WOMAC: 0.89 OHS: 0.87 SF-12 (PCS): 0.62	*Intraclass correlation* WOMAC: 0.88 OHS: 0.87 SF-12 (PCS): 0.58	Better (SF-12 PCS $\Delta = 2.83$, $P < 0.01$) No difference (OHS

(*Continued*)

Table 3. Continued

Author Country/Year	Condition/procedure Recall period Sample size	PROMs	Level of association (correlation coefficient)	Level of agreement	Retrospective health compared to contemporary report[a]
		Lower Extremity Functional Scale (LEFS) Feeling thermometer	SF-12 (MCS): 0.48 LEFS: 0.86 Feeling thermometer: 0.63	SF-12 (MCS): 0.48 LEFS: 0.86 Feeling thermometer: 0.60	$\Delta = -0.04$, $P = 0.96$; SF-12 MCS $\Delta = 2.04$, $P = 0.10$) Worse (WOMAC $\Delta = 2.74$, $P = 0.01$; feeling thermometer $\Delta = -5.06$, $P < 0.01$)
Helfand [17] USA 2010	Urological conditions 6 mo $n = 98$	AUA Symptom Index (SI) Quality of Life (QOL) scores	*Pearson* AUA SI: 0.73 QOL: 0.73	*Kappa* AUA SI: 0.56 QOL: 0.56	Better: AUA SI (recalled mean score 12.2, contemporary 13.1) No difference: QOL (recalled mean score 2.6, contemporary 2.6)

Abbreviations: PROMs, patient-reported outcome measures; SF, Short Form.
[a] Mean difference or proportions different; *P* values.

Correlations tended to be stronger, the shorter the time interval; 1 month or less [14,15] reported strong or very strong agreement. Intervals of 3 months or more resulted in only moderate agreement [12,13]. Another factor associated with the strength of agreement was the type of patient. Most studies that had strong agreement were based on orthopedic patients suggesting patient characteristics or the type of intervention (e.g., elective surgery rather than medical treatment) may influence the relationship.

There was no consistency in the direction of any difference between retrospective and contemporary accounts. One study found that patients tend to recall better baseline health than what they reported in their contemporary PROMs [17], two studies reported the opposite [13,15], one found it varied by PROM [16], and two found no difference [12,14].

The strength of agreement may be limited if the test–retest reliability of the PROM is poor. In Table 4, the reliability estimates for all the measures that were included in studies in Table 3 are presented. Test–retest reliability for all the PROMs used was excellent and higher than the agreements captured when comparing retrospective to contemporary PROMs. This suggests there are additional reasons that influence recall when retrospective PROMs are used.

Table 4. Test–retest reliability of PROMs included in literature review

PROM	Test–retest reliability
SF-12	Physical component: ICC 0.83 [16] Mental component 0.91 [16]
SF-36	ICC = 0.43–0.90 [18]
Oxford Hip Score	Bland Altman coefficient 7.27 [19]
WOMAC	ICC > 0.7 [20]
Lower Extremity Functional Scale	ICC = 0.93 [21]
Feeling thermometer	ICC 0.94 [16]
AUA Symptom Index	$r = 0.92$ [22]
IKDC subjective form	ICC = 0.85–0.99 [23]
ACL-QOL	Standard error of measurement is 6% [24]
WOMET	ICC = 0.79 [14]
KOOS	ICC = 0.75–0.93 [14]

Abbreviations: PROMs, patient-reported outcome measures; SF, Short Form; WOMAC, Western Ontario and McMaster Osteoarthritis Index; AUA, American Urological Association; IKDC, International Knee Documentation Committee; ACL-QOL, Anterior Cruciate Ligament–Quality of Life; WOMET, Western Ontario Meniscal Evaluation Tool; KOOS, Knee Injury and Osteoarthritis Outcome Score; ICC, intraclass correlation coefficient.

3.3. Comparison of retrospective PROMs with population norms

There were 11 studies (Table 5), four from North America [25,26,31,32], four from Australia or New Zealand [29–31,36], and three from Europe [27,33,34]. Eight studies involved fewer than 500 patients (86–472) but three were larger (1,500–3,000 patients). All the studies involved trauma patients apart from one on patients with acute lung injury [31]. Most studies included adults of all ages. The two exceptions were a study of elderly people who had experienced a fractured neck of femur [27] and a study of young adult trauma victims [28].

All reported on a generic PROM: six used a version of the SF (SF-36, SF-12, SF-6); three used the EQ-5D; and two used the Sickness Impact Profile. The time period for retrospective reporting in six studies was less than 1 week [26–31,33]. In the other studies, it extended from a few weeks to 3 months.

Table 5. Studies comparing retrospective PROMs with age–sex standardized general population norms

Author Country/Year	Condition/procedure Recall period Number of patients Patient age and sex	PROMs	Retrospective health compared to general population[a]
Mock [25] USA 2000	Leg injury Weeks (hospital discharge) $n = 302$ Adults (18–64 yrs)	Sickness Impact Profile (SIP)	No difference
Michaels [26] USA 2001	Trauma (blunt force) Days (early in hospital stay) $n = 165$ Adults (mean age 37 yrs); 67% male	SF-36 SIP	No difference
Tidermark [27] Sweden 2002	Fractured neck of femur 12–48 hr after admission $n = 90$ Elderly (mean age 80 yrs)	EQ-5D	No difference
Ameratunga [28] New Zealand 2006	Trauma from motor vehicle accident[b] One day $n = 472$ Young adults (70% 15–44 yrs); 63% male	SF-36	Better than general population No difference from representative sample of drivers
Gabbe [29] Australia 2007	Trauma (mixed) Median 6 days (IQR 3–12 days) $n = 2,388$ Adults	SF-12	Better: SF-12 (physical) mean 50.9 vs. 48.9 ($P < 0.001$) SF-12 (mental) mean 54.5 vs. 52.4 ($P < 0.001$) Differences confined to men and under 55 yrs.
Watson [30] Australia 2007	Trauma (mixed) 4 days (median) $n = 186$ Adults (18–74 yrs)	SF-6D SF-36 Assessment of Quality of Life	Better: Assessment of Quality of Life population norm mean utility 0.83, recalled 0.95 SF-6D population norm mean utility 0.78, recalled 0.92 Better for all age groups ($P < 0.05$).
Gifford [31] USA 2010	Acute lung injury Days–weeks (as soon as patient regained capacity) $n = 136$ Adults (median age 49 yrs; IQR 40–60)	SF-36	Worse: mean paired difference for all SF-36 domains (mean paired differences ranged from 2.6 –17.9) Mean paired difference was significantly better in population norm for all SF-36 domains ($P < 0.01$) except for vitality ($P = 0.12$) Mean retrospective domain scores ranged 56.4–75.6, mean population norm domains scores ranged 58.9–87.6
Lange [32] Canada 2010	Mild traumatic brain injury[c] Median 1.8 months (0.2–8.0) $n = 86$ Adults (mean age 37 yrs; SD 13.7)	British Columbia Post-Concussion Symptom Inventory	Better: overall score ($P < 0.01$) and in 6 of the 13 individual items ($P < 0.05$)
Lyons [33] UK 2011	Trauma (mixed) Within 7 days $n = 1,517$ Adults (median age 37 yrs; IQR 21 –61)	EQ-5D	Better: mean score 3.3% (95% CI 1.9–4.7%) higher
Toien [34] Norway 2011	Trauma (mixed) 17 days (non-ICU) and 44 days (ICU patients) $n = 242$ Adults (mean age 42 yrs)	SF-36	Better: mean score higher ($P < 0.001$)

(*Continued*)

Table 5. Continued

Author Country/Year	Condition/procedure Recall period Number of patients Patient age and sex	PROMs	Retrospective health compared to general population[a]
Wilson [35] New Zealand 2012	Trauma (mixed) 3 mo *n* = 2,856 Adults (18–64 yrs)	EQ-5D	Better: Both the recovered and not recovered groups had significantly better recalled than the population norm Recovered at 5 months: retrospective mean (SD) 0.98 (0.97–0.99) vs. norm 0.85 (0.84–0.86) Not recovered at 5 months: retrospective mean (SD) 0.93 (0.92–0.94) vs. norms 0.85 (0.84–0.87) Recovered at 12 months: retrospective mean (SD) 0.96 (0.96–0.97) vs. norms 0.86 (0.85–0.87) Not recovered at 12 months: retrospective mean (SD) 0.93 (0.93–0.94) vs. norms 0.85 (0.83–0.86)

Abbreviations: PROMs, patient-reported outcome measures; SF, Short Form; EQ-5D, EuroQuol-5D; IQR, interquartile range; SD, standard deviation; CI, confidence interval; ICU, intensive care unit.

[a] Mean difference; *P* value.

[b] Also compared with representative sample of drivers.

[c] Compared with 177 community controls.

All but one study used population norms derived from statutory surveys of the general population. The exception used a matched comparison group drawn from the local community [32]. Also, one study of drivers who had suffered trauma in road accidents was compared not only with population norms but also with a sample of uninjured drivers [28].

Of the 10 studies that used general population norms, six found that patients recalled their health as having been better than the general population [28–30,33–35]. In the four other studies, three found no difference [25–27] and in only one did patients report worse health than the general population [31]. The latter was the only study not focused on trauma patients but on those who had developed acute lung injury who were likely to have been in a poor state of health before being hospitalized. The two studies that compared patients with matched samples rather than the general population reported either no difference [28] or better recalled health [32].

4. Discussion

4.1. Comparison of retrospective and contemporary PROMs

Only six studies have compared retrospective and contemporary PROMs. Although the majority of the comparisons (21 of 30) revealed a strong or very strong association (correlation coefficients of over 0.68), the rest were moderate. Levels of agreement for continuous measures were more consistent with 20 of 24 comparisons being strong or very strong. In contrast, comparisons of categorical measures showed only fair to moderate agreement. Stronger associations were observed for indices (than for individual items), for shorter time periods (1 month or less), and for elective surgery patients than for those with medical conditions or treatments. The direction of differences between retrospective and contemporary PROMs also showed no consistent pattern and appeared to be dependent partly on the PROM being used.

Retrospective PROMs may be influenced for three reasons: recall bias; response shift; and lack of validity of the PROM. Recall bias arises because details may go unnoticed and never be stored; new information may be added to stored memories altering the details; and over time, events may be systematically distorted [6]. Recall is influenced by the time interval between the event and the time of its assessment: the longer the interval, the higher the probability of recall bias [37]: 20% of details of an event have been found to be irretrievable after 1 year and 50% are irretrievable after 5 years [38].

Response shift refers to the change in perception that can occur when circumstances change [39,40]. For example, a patient's perception of the severity of a disability or their quality of life may change following treatment. This tends to diminish the assessment of pretreatment severity and thus underestimate the benefits of

the treatment. An example of this is when the term "severe," has a different meaning for the same person in one occasion compared with a previous occasion due to new experiences. This is known as recalibration. Moreover, subjective values may also change over time so that physical, social and psychological aspects of HRQL may be prioritized differently after certain experiences, known as reprioritization. Patients may also redefine the construct in question and attribute new meanings to it, known as scale reconceptualization [41].

It is possible that the validity of PROMs will be jeopardized when determining retrospective health if the recall interval is lengthy. Most PROMs have been validated for the recall of a person's health over the recent past (between 1 day to past 4 weeks). Indeed, many PROMs are based on patients' reports of their health over the preceding few weeks. However, if patients are required to recall their health for longer periods, the validity of the instrument cannot be assumed.

For comparisons of healthcare providers or over time, recall bias and response shift will only matter if there is a systematic difference in behavior between groups of patients being compared (e.g., patients attending different hospitals). There is no evidence that such differences exist within countries though some differences have been demonstrated between countries [42].

4.2. Comparisons of retrospective PROMs and population norms

The studies comparing retrospective PROMs with population norms were inevitably limited to generic instruments because disease-specific PROMs are rarely collected in general population surveys and hence limits the availability of population data to generic PROMs. The generalizability of the findings is further limited by the focus of all but one study on trauma victims. The finding that most studies observed that trauma patients recalled their preinjury health as better than average may reflect that patients (mostly car drivers) are fitter and healthier than the general population [19]. Although response shift may have contributed, the likelihood that trauma patients were healthier is supported by evidence that rates of sports injuries and gunshot wounds are higher in fitter members of the population [29−31,33]. This difference is further exaggerated as national population norms are derived from household surveys that include institutionalized individuals. In contrast, the one study of elderly people experiencing a stress fracture related to poor bone density found no difference from the general population (age−sex standardized) [27]. This is also consistent with the one study in which patients recalled worse health than the general population which focused on acute lung injury [31].

There may be a case for the purposes of estimating pre-event health status that estimates could be adjusted for the presence of long-term conditions to reduce overestimation. The findings also suggest the potential of underestimating the prior health of patients if population norms are used directly as surrogates in cases where the patient population involved are younger adults. However, this underestimation may be small and may mostly affect studies in this specific cohort of patients.

4.3. Limitations

There are several limitations to consider. First, only one author (E.K.) carried out the search, paper selection, and quality appraisal. Although uncertainties were discussed and resolved with the other author, the reliability of the review would have been enhanced by double-reviewing. Second, comparisons of retrospective and contemporary PROMs that have been studied are dominated by orthopedic surgery (four of six studies) and by studies in North America (four of six). Thus, the generalizability of the findings must be treated with caution. Third, many of the studies that investigated retrospective recall were too small to perform subgroup analysis to take into account of clinical characteristics such as severity of illness. Finally, the generalizability of the comparisons of retrospective PROMs and population norms are even more limited with 10 of the 11 studies focused on trauma patients. In addition, only generic PROMs were considered, but this is understandable given that population norms are not available for disease-specific PROMs.

4.4. Implications for policy and research

Making judgments as to which of contemporary and retrospective reports are the more valid is unclear. Contemporary reports are usually considered the "gold standard" so if retrospective reports differ, it is the latter that are judged to be "unreliable." However, in the context of PROMs, from a patient's point of view, the way they recall their previous health may be of greater relevance to them and to assessing the quality of health care than how patients actually assessed it at the time. In this situation, the retrospective report could be viewed as the "gold standard." Rather than attach different values to the two types of PROM (in other words, judging whether contemporaneous collection is more or less valid than recalled collection), it is best just to consider the extent to which they differ and the implications both for the use of PROMs in clinical management and in provider comparisons. As long as data are collected in the same way in different providers, then comparisons will not be undermined.

Our knowledge of the use of retrospective PROMs in the United Kingdom is extremely limited: the relevance of findings in other countries is uncertain given the potential influence of culture and other contextual factors; existing studies of unexpected emergency admissions are limited largely to trauma care; and there have been no published

attempts to study both of the issues addressed in this review in a combined study (i.e., retrospective vs. contemporary vs. population norms). Until further research has been conducted, the best policy for using PROMs in emergency admissions will remain uncertain.

The key methodological challenges that require further research are as follows: detailed investigation of the relationship between retrospective and contemporary PROMs (inevitably in elective conditions) which should also explore the influence of patient characteristics and of methodological factors on the relationship; determination of the potential use of population norms as a low-cost alternative to retrospective PROMs; and testing the feasibility of retrospective PROMs and population norms in a variety of unexpected emergency hospital admissions.

References

[1] Black N. Patient reported outcome measures could help transform healthcare. BMJ 2013;346:f167.

[2] Health & Social Care Information Centre. Hospital Episode Statistics 2013-14. Available at http://www.hscic.gov.uk/hes. Accessed October 7, 2015.

[3] Hennekens CH, Buring JE. Epidemiology in medicine. Philadelphia: Lippincott Williams & Wilkins; 1987.

[4] McKenzie E. Measuring disability and quality of life postinjury. In: Rivara FP, Cummings P, Koepsell PD, Grossman DC, Maier RV, editors. Injury control: a guide to research and program evaluation. Cambridge, UK: Cambridge University Press; 2001.

[5] Stull DE, Leidy NK, Parasuraman B, Chassany O. Optimal recall periods for patient-reported outcomes: challenges and potential solutions. Curr Med Res Opin 2009;25(4):929–42.

[6] Schmier JK, Halpern MT. Patient recall and recall bias of health state and health status. Expert Rev Pharmacoecon Outcomes Res 2004; 4(2):159–63.

[7] Damiano AM, Pastores GM, Ware JE Jr. The health-related quality of life of adults with Gaucher's disease receiving enzyme replacement therapy: results from a retrospective study. Qual Life Res 1998;7: 373–86.

[8] Guyatt GH, Norman GR, Juniper EF, Griffith LE. A critical look at transition ratings. J Clin Epidemiol 2002;55:900–8.

[9] Lucas NP, Macaskill P, Irwig I, Bogduk N. The development of a quality appraisal tool for studies of diagnostic reliability. J Clin Epidemiol 2010;63:854–61.

[10] Taylor R. Interpretation of the correlation coefficient: a basic review. J Diagn Med Sonography 1990;6:35–9.

[11] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.

[12] Emberton M, Challands A, Styles RA, Wightman JA, Black N. Recollected versus contemporary patient reports of pre-operative symptoms in men undergoing transurethral prostatic resection for benign disease. J Clin Epidemiol 1995;48:749–56.

[13] Lingard EA, Wright EA, Sledge CB. Pitfalls of using patient recall to derive preoperative status in outcome studies of total knee arthroplasty. J Bone Joint Surg Am 2001;83-A:1149–56.

[14] Bryant D, Norman G, Stratford P, Marx RG, Walter SD, Guyatt G, et al. Patients undergoing knee surgery provided accurate ratings of preoperative quality of life and function 2 weeks after surgery. J Clin Epidemiol 2006;59:984–93.

[15] Howell J, Xu M, Duncan CP, Masri BA, Garbuz DS. A comparison between patient recall and concurrent measurement of preoperative quality of life outcome in total hip arthroplasty. J Arthroplasty 2008; 23:843–9.

[16] Marsh J, Bryant D, MacDonald SJ. Older patients can accurately recall their preoperative health status six weeks following total hip arthroplasty. [Internet]. J Bone Joint Surg Am 2009;91: 2827–37.

[17] Helfand BT, Fought A, Manvar AM, McVary KT. Determining the utility of recalled lower urinary tract symptoms. Urology 2010;76: 442–7.

[18] Ware JEJ, Snow KK, Kosinski M, Gandek B. SF-36 Health Survey: manual and interpretation guide. Boston: The Health Institute, New England Medical Centre; 1993.

[19] Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. J Bone Joint Surg Br 1996;78:185–90.

[20] Stucki G, Sangha O, Stucki S, Michel BA, Tyndall A, Dick W, et al. Comparison of the WOMAC (Western Ontario and McMaster Universities) osteoarthritis index and a self-report format of the self-administered Lequesne-Algofunctional index in patients with knee and hip osteoarthritis. Osteoarthritis Cartilage 1998;6: 79–86.

[21] Smith S, Cano S, Lamping D, Staniszewska S, Browne J, Lewsey J, et al. Patient-Reported Outcome Measures (PROMs) for routine use in Treatment Centres: recommendations based on a review of the scientific evidence. Report to the Department of Health, 2005. Available at https://www.lshtm.ac.uk/php/departmentofhealthservicesresearchandpolicy/assets/promsnickblack2005.pdf

[22] Barry MJ, Fowler FJ Jr, O'Leary MP, Bruskewitz RC, Holtgrewe HL, Mebust WK, et al. The American Urological Association symptom index for benign prostatic hyperplasia. J Urol 1992;148:1549–57.

[23] Kanakamedala AC, Anderson AF, Irrgang JJ. IKDC Subjective Knee Form and Marx Activity Rating Scale are suitable to evaluate all orthopaedic sports medicine knee conditions: a systematic review. Joint Disord Orthopaedic Sports Med 2016;1(2). http://dx.doi.org/10.1136/jisakos-2015-000014.

[24] Mohtadi N. Development and validation of the quality of life outcome measure (questionnaire) for chronic anterior cruciate ligament deficiency. Am J Sports Med 1998;26:350–9.

[25] Mock C, MacKenzie E, Jurkovich G, Burgess A, Cushing B, deLateur B, et al. Determinants of disability after lower extremity fracture. J Trauma 2000;49:1002–11.

[26] Michaels J, Madey SM, Krieg JC, Long WB. Traditional injury scoring underestimates the relative consequences of orthopedic injury. J Trauma 2001;50:389–95.

[27] Tidermark J, Zethraeus N, Svensson O, Törnkvist H, Ponzer S. Femoral neck fractures in the elderly: functional outcome and quality of life according to EuroQol. Qual Life Res 2002;11:473–81.

[28] Ameratunga SN, Norton RN, Connor JL, Robinson E, Civil I, Coverdale J, et al. A population-based cohort study of longer-term changes in health of car drivers involved in serious crashes. Ann Emerg Med 2006;48(6):729–36.

[29] Gabbe BJ, Cameron PA, Graves SE, Williamson OD, Edwards ER. Preinjury status: are orthopaedic trauma patients different than the general population? J Orthop Trauma 2007;21:223–8.

[30] Watson WL, Ozanne-Smith J, Richardson J. Retrospective baseline measurement of self-reported health status and health-related quality of life versus population norms in the evaluation of post-injury losses. Inj Prev 2007;13:45–50.

[31] Gifford JM, Husain N, Dinglas VD, Colantuoni E, Needham DM. Baseline quality of life before intensive care: a comparison of patient versus proxy responses. Crit Care Med 2010;38(3): 855–60.

[32] Lange RT, Iverson GL, Rose A. Post-concussion symptom reporting and the "good-old-days" bias following mild traumatic brain injury. Arch Clin Neuropsychol 2010;25(5):442–50.

[33] Lyons R, Kendrick D, Towner EM, Christie N, Macey S, Coupland C, et al. Measuring the population burden of injuries—implications for global and national estimates: a multi-centre prospective UK longitudinal study. PLoS Med 2011;8(12):e1001140.
[34] Tøien K, Bredal IS, Skogstad L, Myhren H, Ekeberg O. Health related quality of life in trauma patients. Data from a one-year follow up study compared with the general population. Scand J Trauma Resusc Emerg Med 2011;19(1):22.
[35] Wilson R, Derrett S, Hansen P, Langley J. Retrospective evaluation versus population norms for the measurement of baseline health status. Health Qual Life Outcomes 2012;10(1):68.
[36] Basso O, Olsen J, Bisanti LKW. The performance of several indicators in detecting recall bias. Epidemiology 1997;8:269−74.
[37] Skowronski JJ, Betz AL, Thompson CP, Shannon L. Social memory in every day life: recall of self-events and other-events. J Pers Soc Psychol 1991;60:831−43.
[38] Bradburn N, Rips L, Shevell S. Answering autobiographical questions: the impact of memory and inference on surveys. Science 1987;236:157−61.
[39] Sprangers M. Response-shift bias: a challenge to the assessment of patients' quality of life in cancer clinical trials. Cancer Treat Rev 1996;22(Suppl A):55−62.
[40] Visser M, Oort F, Sprangers M. Methods to detect response shift in quality of life data: a convergent validity study. Qual Life Res 2005;14:629−39.
[41] Howard JS, Mattacola CG, Howell DM, Lattermann C. Response shift theory: an application for health-related quality of life in rehabilitation research and practice. J Allied Health 2011;40:31−8.
[42] Black NA, Glickman ME, Ding J, Flood AB. International variation in intervention rates: what are the implications for patient selection? Int J Tech Ass Health Care 1995;11:719−35.

Chapter 3

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Esther Kwong
Principal Supervisor	Nick Black
Thesis Title	The use of patient reported outcome measures (PROMs) for evaluating emergency admissions

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?	Quality of Life Research		
When was the work published?	2018		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	Please refer to details on the following page

Student Signature: _____ **Date:** 24 August 2018

Supervisor Signature: _____ **Date:** 24 August 2018

3.1 Chapter 3

Agreement between retrospectively and contemporaneously collected patient-reported outcome measures (PROMs) in hip and knee replacement patients

Chapter 3 compares retrospective with contemporary PROMs in elective admissions to investigate the agreement between contemporary and retrospective PROMs in the English NHS context. This work was conducted in elective orthopaedic patients in a longitudinal cohort study as contemporary PROMs are already collected from Orthopaedic patients undergoing elective hip and knee arthroplasty surgery. Patients who were participating on the NHS National PROMs Programme were recruited to the study during their inpatient period following their elective surgery. Consented patients then completed a retrospective PROM, where they were asked postoperatively about how they were before their hospital admission. Their retrospective and contemporary PROMs questionnaires were linked to compare the agreement between the scores.

I was the chief investigator for this study; I led the study design, developed the study protocol with guidance from Professor Nick Black, and consulted with orthopaedic clinical leads at the local sites, Mr Joyti Saksena, Mr Mahbub Alam, Mr Rej Bumbra, and Professor Fares Haddad from the participating hospital sites. I consulted and piloted the retrospective PROMs questionnaires designs with the CLARHC North Thames public and patients' panel, and revised the layout for ease of use from the feedback and comments received.

I held regular meetings with orthopaedic PROMs co-ordinators at the study sites (Miss Jamila Kassam, Mrs Ameena Hare, Miss Hazera Mahdiya and Mrs Ursula Knight) to understand local data collection pathways and embedded the study into established pathways in the study design. I conducted training for all local sites data collection teams as although staff were familiar with PROMs collection as part of the national PROMs programme (contemporary pre-operative Q1), collection of a retrospective PROM during the inpatient stay

post-operatively is novel for the teams. I visited sites regularly during data collection to improve and support smooth running of the study. I was able to gain an understanding of designing and leading a cohort study for collecting retrospective PROMs during an inpatient hospital admission in acute hospital trusts. These methods subsequently informed the study designs for the feasibility studies of PROMs for emergency patient cohorts.

I conducted the deterministic linkage of PROMs data, and was responsible for the statistical analysis with advice from Dr Jenny Neuburger. I prepared a first draft of the manuscript. All co-authors made comments on successive drafts and approved the final version before journal submission. I acted as a guarantor of the final published version.

Quality of Life Research
https://doi.org/10.1007/s11136-018-1823-6

Agreement between retrospectively and contemporaneously collected patient-reported outcome measures (PROMs) in hip and knee replacement patients

Esther Kwong[1] · Jenny Neuburger[2] · Nick Black[1]

Accepted: 21 February 2018

Abstract

Purpose To investigate the relationship between retrospectively and contemporaneously collected patient-reported outcome measures (PROMs) and the influence on this relationship of patients' age and socio-economic status and the length of time.

Methods Patients undergoing hip or knee replacement in four hospitals who had completed a pre-operative questionnaire were invited to recall their pre-operative health status shortly after surgery. The questionnaires included a disease-specific (Oxford Hip Score; Oxford Knee Score) and generic (EQ-5D-3L) PROM. Consistency and absolute agreement between contemporary and retrospective reports were investigated using intraclass correlations (ICCs). Differences were visualised using Bland–Altman plots. Linear regression analysis explored whether retrospective can predict contemporary PROMs.

Results Patients' recalled health statuses were similar to their contemporaneous reports, with no significant systematic bias. Absolute agreement for disease-specific PROMs was very strong (ICC 0.82) and stronger than for the generic PROM (ICC 0.60, 0.62). Agreement was consistently strong across the range of severity of a patient's condition, age and socio-economic status. Patients' age and socio-economic status had no significant influence on size of difference and direction of recall, although reliability of recall was slightly worse among the over-75s versus under-60s for hips (Oxford Hip Score ICC 0.88 vs. 0.78). Mean retrospective PROMs for groups or populations of patients can reliably predict what mean contemporary reports of PROMs would have been.

Conclusion Retrospective PROMs can be used to obtain a baseline assessment of health status when contemporary collection is not feasible or cost effective. Research is needed to determine the feasibility of retrospective PROMs in emergency admissions.

Keywords Patient-reported outcome measures · Health status · Health-related quality of life · Retrospective · Recall · Agreement

Introduction

Patient-reported outcome measures (PROMs) have the potential to transform health care delivery through enhancing the clinical management of patients and assessing the quality of providers' performance [1, 2]. To date, the use of PROMs in assessing the outcome of hospital admissions has inevitably been restricted to elective surgery in which before and after measurements of patients' symptoms, functional status and health-related quality of life can be compared. The most ambitious example of this covers four elective surgical procedures in the NHS in England [3].

The challenge of using PROMs for emergency admissions, which account for 40% of hospital inpatients in England, has not been addressed and yet this is an area of

✉ Esther Kwong
esther.kwong@lshtm.ac.uk

Jenny Neuburger
jenny.neuburger@lshtm.ac.uk

Nick Black
nick.black@lshtm.ac.uk

[1] Department of Health Services Research, Faculty of Public Health and Policy, London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, UK

[2] London School of Hygiene and Tropical Medicine, London, UK

Published online: 26 February 2018

increasing resource use, political importance and concern about variations in quality of care [4]. The methodological challenge is how to quantify outcome when a patient's health status before their sudden and unexpected ill-health that led to an emergency hospital admission is, inevitably, not available. One potential solution would be if patients were able to recall accurately their health status before the admission. If they could, then a retrospective (or recalled) PROM would offer a means of obtaining their baseline health status in the absence of a prospectively collected contemporary report.

A recent literature review on the relationship between retrospective and contemporary health status reports found strong agreement when the recall period is short [5]. However, only six studies have been undertaken of which only one was conducted in the UK [6]. The relevance of findings from other countries is uncertain given the potential influence of culture and other contextual factors. In addition, only two studies considered the influence of patients' characteristics, such as social demographic factors, on the relationship. Both studies found that agreement was slightly weaker in older patients [7, 8].

Our aim was to investigate the relationship between retrospective and contemporary PROMs in England (inevitably, in elective conditions) and to explore the influence on the relationship of two patient characteristics (age, socio-economic status) and the length of time between the two data collection points. Contemporary reports are often considered the 'gold standard' so if retrospective reports differ, it is the latter that are judged 'unreliable'. However, in the context of PROMs, from a patient's point of view the way they recall their previous health may be of greater relevance to them and to assessing the quality of health care than how they assessed it at the time. Rather than assuming one as the 'gold standard' over the other type of PROM, we consider the extent to which they agreed. We hypothesise that if the two agree then one can substitute for the other without any impact on assessment of the impact of health care interventions. If they differ, it would be necessary to consider the reasons for this and its implications for the use of PROMs in clinical management and in provider comparisons in emergency admissions.

Methods

Sample

This is a multi-centre study of patients undergoing either hip or knee arthroplasty (primary operation or revision surgery) in four hospitals, which were part of the North Thames Academic Health Science Network (UCL Partners), and CLAHRC. Health Research Authority ethics approval was obtained from North East – Newcastle & North Tyneside 2 Research Ethics Committee (REC Ref: 16/NE/0081).

Patients were eligible if, as part of the National PROMs Programme, they had completed a PROM questionnaire before undergoing surgery (Q1), either at a pre-operative assessment clinic or on their day of admission. They were invited to complete a retrospective PROM questionnaire (QR) in the immediate post-operative period prior to discharge asking them to recall their health status during the 4 weeks prior to surgery. Written informed consent was obtained.

Patients' QR was deterministically linked to their contemporaneous PROMs data (Q1) using a hierarchy of patient identifiers: NHS number, date of birth, postcode and date of birth and postcode combined.

Questionnaires

The self-reported questionnaires included socio-demographic information: age; sex; living arrangement (with family/friends, alone, other). Socio-economic status (SES) was measured with national quintiles of the Index of Multiple Deprivation based on patients' residential postcode [9]. Self-reported health included co-morbidities (from a list of 12 conditions); duration of primary condition (< 1, 1–5, 6–10, > 10 years); primary or revision surgery; disease-specific PROM (Oxford Hip Score or Oxford Knee Score); and a generic PROM (EQ-5D-3L)—the latter was used as it was the version used in the National PROMs Programme in England for elective surgery at the time.

The Oxford Hip Score (OHS) is a disease-specific PROM for patients undergoing total hip replacement to capture symptoms and functional status [10]. It has good face validity, construct validity and reliability, and is sensitive to change. The Oxford Knee Score (OKS) is the knee arthroplasty equivalent [11]. For both PROMs, respondents answer 12 questions to assess pain and mobility in relation to the relevant joint. Each item can be scored from 0 (severe problem) to 4 (no problem). Summated scores provide an overview, from 0 (worst) to 48 (best) health statuses [12].

For the Oxford Scores, instructions were adapted to enable usage for retrospective assessment (QR) by including a statement on the timeframe with the following wording; 'We are interested in finding out about the problems you have had with the hip (knee) on which you have had surgery. Please let us know how you were before your operation'. This kept the wording similar to the instructions for the prospective version use in the National PROMs programme (Q1); 'We are interested in finding out about the problems you have had with the hip (knee) on which you are about to have surgery'. The tense of individual questions were also altered, e.g. Q1: 'During the past 4 weeks…How would you describe the pain you usually have from your knee?' was changed to 'During

the past 4 weeks before your operation…How would you describe the pain you usually had from your knee?'.

The EQ-5D-3L has five questions that investigate the domains of mobility, usual activities, self-care, pain/discomfort and anxiety/depression [13]. For each of these questions, the respondent chooses from three responses indicating the level of their function. A multi-attribute utility score where death and perfect health are represented by 0 and 1 are calculated [14]. Scores less than 0 are considered worse than death and 1 is the maximum score possible. The EQ-VAS (a visual analogue scale) was also included in the questionnaires but this was not included in the analysis of the results, due to missing data and respondents not completing it according to instructions [15].

For the EQ-5D-3L, wording was adapted to provide instructions suitable for retrospective assessment with 'before your operation' in place of 'today'. The full instructions on QR were: 'By placing a tick in one box in each group below, please indicate which statements best describe your own health state before your operation'. Each statement of individual items was changed to past tense (e.g. 'I have no problems walking about' was changed to 'I had no problems walking about)'.

Sample size

Sample size was designed to achieve the required degree of precision in the estimation of the ICC. For example, a sample of 200 patients would give a two-sided 95% confidence interval of 0.14 if the ICC was 0.7 (ICC CI 0.62–0.76). Consequently, we selected a total sample of 400 (200 for each procedure), which meant that the width of the CI (0.14) was less than the width of bands used to define categories of agreement (see below). It also provided sufficient statistical power for some sub-group analyses [16, 17].

Statistical analysis

Agreement between patients' retrospective and contemporaneous PROMs scores was judged both in terms of absolute agreement and consistency. It was assumed that both time points measure the same construct and should thus be in strong absolute agreement. However, while any systematic differences in recall could reduce absolute agreement, if patients retained their Q1 and QR ranking order, then there would still be consistency in the scores. We therefore also looked at consistency which could be useful from a policy perspective as even if scores lacked absolute agreement but remained consistent, then PROMs retrospective scores would be useful in assessing provider performance. Agreement was categorised as 0–0.20 weak, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 strong and 0.81–1 very strong [18].

We calculated separate intraclass correlations for absolute agreement (ICC(A,1)) and consistency (ICC(C,1)) using the definitions given by McGraw and Wong [16], as well as Pearson's correlation coefficient as a measure of association. The analysis was conducted using Stata version 14 [17]. The ICCs were calculated using repeated measures of analysis of variance (ANOVA) which divides the variance into three components: between-subjects (patients), within-subjects (contemporaneous recall) and error. They are presented with their 95% confidence intervals.

To explore patterns of differences in the contemporary and retrospective score visually, we used a version of the Bland–Altman plot that accounts for trend. Individual differences in scores were plotted against the mean of the two scores, and a regression model was used to calculate the limits of agreement [19]. As neither the contemporaneous nor the retrospective method is assumed to be a gold standard, the mean of the two is the best estimate of the true health status and most appropriate for the *x*-axis [20].

Finally, linear regression analysis was conducted to explore whether a patient's retrospective PROM is able to predict their contemporary PROM, judged from differences in their predicted (based on retrospective) and contemporary PROM (mean absolute error). Scatterplots of contemporary score (*y*-axis) against the retrospective score (*x*-axis) are shown in Fig. 1, along with the mean predicted score (linear fit) and 95% confidence intervals. The wider lines show the 95% confidence intervals around individual predictions, taking into account the residual variation in individual scores.

The influence on the relationship between retrospective and contemporary PROMs of two patient characteristics (age and social-economic status) and one logistical (length of time between the two data collection points) was explored using linear regression analysis; ICCs were also calculated for age subgroups.

Results

Patient characteristics

The required sample size of 400 in total was exceeded. Of the 406 hip replacement patients who had completed a Q1 and were invited to complete a QR, 244 (60%) did so. Equivalent figures for knee replacement were 276 out of 486 (57%). It was not possible to link data from the two questionnaires for some patients (20 hip; 16 knee) and the disease-specific PROM was not fully completed by some patients (20 hip; 21 knee) (Appendix). This left 204 hip and 239 knee patients for the analysis.

The sample was broadly similar to the population of patients completing pre-operative PROM questionnaires in 2009–2010, the latest year for which published data exist

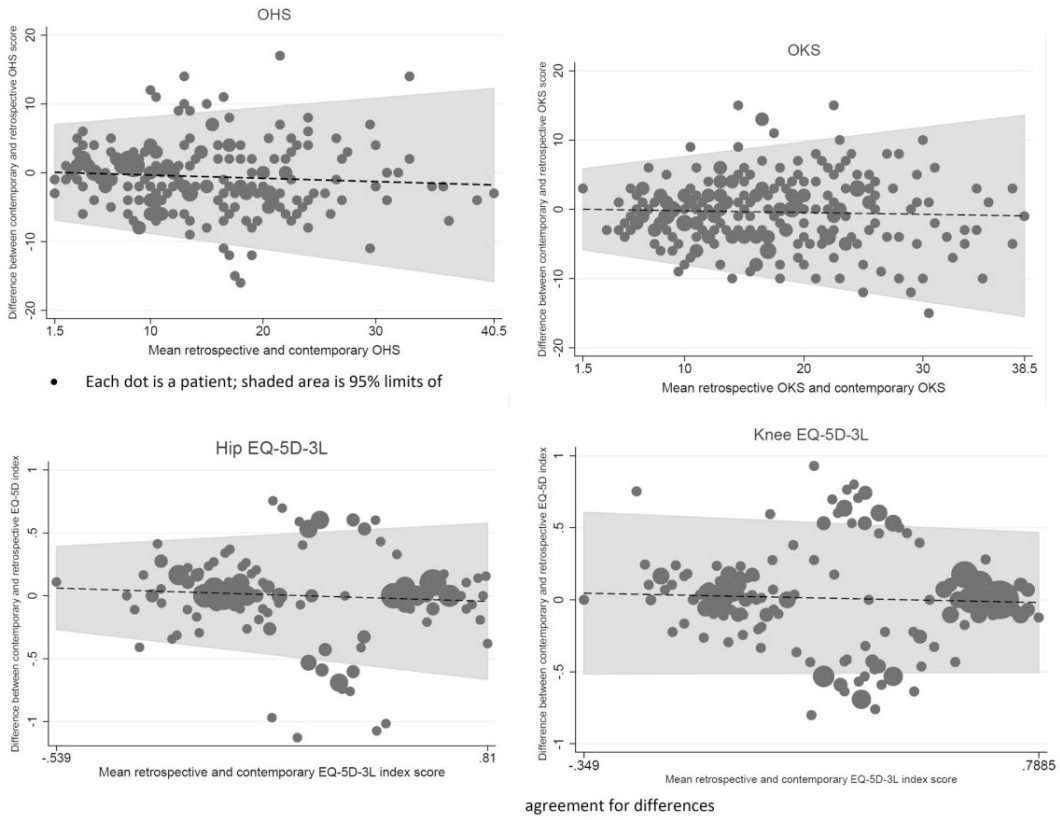


Fig. 1 Patterns of differences in contemporary and retrospective PROMs (OHS, OKS and EQ5D) adjusting for trend. Each dot is a patient; shaded area is 95% limits of agreement for differences

[21, 22]. There were some small differences (Table 1). The hip replacement sample was slightly older (mean age 69.1 vs. 67.7 years) and more likely to be female (67 vs. 61%), and to live alone (34 vs. 28%). The knee patients were also more likely to live alone (29 vs. 25%). For both operations, patients reported having more severe conditions (mean OHS 15.1 vs. 18.2; mean OKS 17.4 vs. 19.3; knee symptoms for over 5 years 55 vs. 44%). This may reflect selection bias in the sample or a change between 2009/2010 and 2016 in the severity of patients' conditions.

While most patients (75%) completed their QR within 50 days of having completed the contemporary Q1, for 3% it was over 3 months (due to delays in surgery following their pre-operative assessment). The median length of time was 30 days (IQR 14–54 days).

Comparison between retrospective and contemporary PROMs

The mean difference between retrospective and contemporary scores was small for all PROMs and both operations (Table 2). The direction of the difference was consistent: patients reported slightly lower scores (worse health) in the retrospective questionnaire compared to the contemporary reports. However, none of the differences were statistically significant.

Absolute agreement and consistency were very strong for both disease-specific PROMs. Agreement on the EQ-5D-3L was also strong, although weaker than for the disease-specific PROM. The level of agreement was consistent across the range of severity of pre-operative health (i.e. there was little systematic bias) as shown by the flat trend lines (Fig. 2). The clustering seen for the EQ-5D-3L results

Table 1 Characteristics of samples compared with population of patients (2009–2010) [21, 22]

Characteristic	Hip replacement		Knee replacement	
	Sample	Population	Sample	Population
Sex: female (%)	136 (67)	(61)	152 (61)	(57)
Age—mean (SD)	69.1 (12.4)	67.7	68.7 (9.0)	68.7
Living arrangements (%)				
With family/friends	65	72	70	75
Alone	34	28	29	25
Other	1	0	1	0
Duration of symptoms (%)				
0–5 year	79	81	45	57
> 5 years	21	19	55	44
Primary operation (%)	91	90	90	92
Disease-specific PROM Q1—mean (SD)	15.1 (8.7)	18.2	17.4 (8.2)	19.3
EQ-5D-3L Q1—mean (SD)	0.24 (0.33)	0.36	0.35 (0.32)	0.43

Table 2 Agreement between contemporary and retrospective PROMs

PROM	Mean Q1:QR	Mean difference (95% CI)	*p* value	ICC absolute agreement (95% CI)	ICC consistency (95% CI)
OHS	15.07:14.56	0.51 (−0.19 to 1.23)	0.15	0.82 (0.77–0.86)	0.82 (0.77–0.86)
OKS	17.36:17.02	0.34 (−0.30 to 0.98)	0.29	0.82 (0.77–0.85)	0.82 (0.77–0.85)
Hip EQ-5D	0.24:0.22	0.02 (−0.02 to 0.06)	0.3	0.62 (0.53–0.69)	0.62 (0.56–0.69)
Knee EQ-5D	0.35:0.32	0.03(−0.01 to 0.07)	0.16	0.60 (0.51–0.67)	0.60 (0.51–0.67)

from there being only three possible levels of response to each item and the way one dimension, pain/discomfort, is weighted heavily in the index score. Therefore, patients who shifted in their level in the pain dimension resulted in a more marked change in their index score, while the average of their two scores was in the middle (see Fig. 1 EQ-5D). In contrast, there was greater concordance between retrospective and contemporary scores in patients who reported either no or extreme pain/discomfort and who did not shift their responses (with their average of their two score remaining at one extreme or the other, i.e. responses seen in the clusters to the most left and furthest right on the horizontal axis).

Prediction of contemporary using retrospective PROMs scores

Patients' retrospective PROMs were able to predict contemporary scores for all three PROMs. The mean absolute error for the prediction model were 3.89 (Q1 SD 8.7) and 3.86(Q1 SD 8.2) for the Oxford Hip and Knee scores and 0.20 and 0.21 for generic EQ-5D scores at the individual level (Table 3). At the group level, this would translate into an even smaller error. The 95% confidence intervals for the mean predicted score (group prediction) is extremely narrow (Fig. 2).

Influences on relationship between contemporary and retrospective PROMs

Agreement between the retrospective and contemporary PROM was strong or very strong across the age range, although slightly weaker with increasing age. For hip patients, the ICC declined from 0.88 for those aged 60 years or younger to 0.78 for those over 75 years (p value < 0.05). The difference for knee patients was less (ICC 0.80 vs. 0.78). There was no evidence of any systematic differences in the magnitude and the direction of recall with patients' age as well as socio-economic status for both Oxford Hip and Oxford Knee Scores. There was some evidence of a slight systematic difference with patients' age on EQ-5D-3L for knee patients (Table 4).

The difference in mean contemporary and retrospective scores was not associated with the time interval between Q1 and QR. The difference in Oxford Knee Score decreased by 0.013 (95% CI − 0.03 to 0.007) and knee EQ-5D-3L score decreased 0.0003 (95% CI − 0.001 to 0.0007). The difference for Oxford Hip Score increased by 0.006 (95% CI − 0.01 to 0.02) per day, and the hip EQ-5D-3L score increased by 0.0001 (− 0.0009 to 0.001) per additional day.

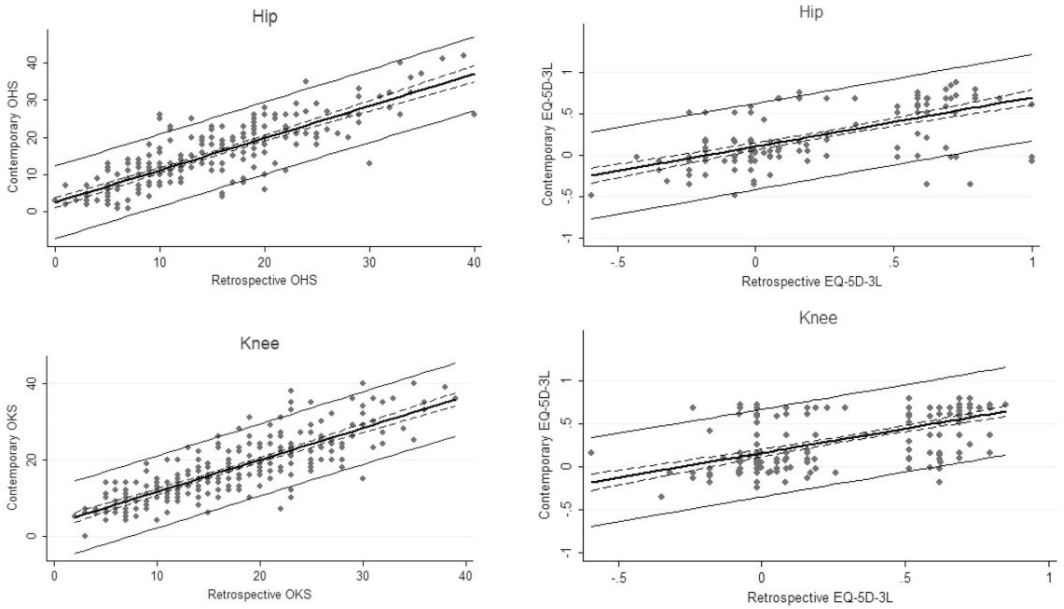


Fig. 2 Contemporary PROM by retrospective PROM linear regression with 95% intervals for individual (solid line) and group (dotted line) contemporary PROMs predictions. Dots represent actual PROM scores, and the solid line the predicted contemporary PROMs scores with 95% intervals for individual and group predictions

Table 3 Retrospective scores as a predictor of contemporary PROMs

PROM	Pearson correlation coefficient (*r*)	Mean absolute error	Coefficient B (95% CI)
Oxford Hip Score	0.82	3.89	0.85 (0.77–0.94)
Oxford Knee Score	0.82	3.68	0.84 (0.76–0.91)
Hip EQ-5D	0.58	0.20	0.62 (0.50–0.73)
Knee EQ-5D	0.56	0.21	0.59 (0.48–0.70)

Discussion

Main findings

In representative samples of patients undergoing elective hip or knee replacement, their retrospective assessment of their pre-operative health status was similar to their contemporaneous reports. Although patients tended to recall their health as being slightly worse than reported at the time across all measures, the differences were small and none was statistically significant. This could result in a slightly higher estimation of the benefits of surgery. The level of agreement between contemporary and recalled PROM scores was very strong for the disease-specific ones, and strong for the generic PROM.

The strength of agreement was consistent regardless of the severity of a patient's primary condition. In addition, two social characteristics of patients, their age and their socio-economic status, had little or no significant influence on the relationship between retrospective and contemporary reports. It was also apparent that mean retrospective PROMs for groups of patients could reliably predict what mean contemporary reports of PROMs would have been.

Comparison to existing studies

These results confirm the findings of the four published studies which also found strong and very strong agreement between retrospective and contemporary PROMs which used continuous rather than categorical data [8, 23–25]. These previous studies also found that agreement

Table 4 Mean difference and adjusted mean difference between retrospective and contemporary PROMs by patients' socio-economic status (SES) and age

Patient characteristic	Hip replacement (OHS)					Knee replacement (OKS)				
	n (%)	Mean Q1:QR	Mean difference	Adjusted difference	*p* value	*n* (%)	Mean Q1-QR	Mean difference	Adjusted difference	*p* value
SES (quintiles)				*Adjusted by age					*Adjusted by age	
1 (Least deprived)	15 (7.4)	15.8–17.0	−1.13	−0.72 (−3.64 to 2.20)		26 (10.6)	18.1–20.5	−2.35	−1.74 (−4.18 to 0.65)	
2	36 (17.7)	15.1–13.3	1.89	2.14 (0.01 to 4.29)		43 (17.6)	19.4–18.3	1.02	1.46 (−0.60 to 3.53)	
3	57 (27.9)	14.8–14.9	−0.14	Reference group	0.16	48 (19.6)	16.1–16.5	−0.31	Reference group	0.10
4	54 (26.5)	16.4–15.2	1.25	1.44 (−0.44 to 3.34)		69 (28.2)	17.4–16.9	0.52	0.85 (−1.92 to 1.75)	
5 (Most deprived)	42 (20.6)	13.5–13.4	0.14	0.26 (−1.76 to 2.28)		58 (23.7)	14.7–15.1	−0.36	−0.81 (−2.73 to 1.11)	
Age (years)				*Adjusted by SES					*Adjusted by SES	
≤60	49 (24)	16.3–15.8	0.53	−0.27 (−2.09 to 1.53)		43 (18)	16.9–15.3	1.62	2.00 (0.29 to 3.73)	
61–75	85 (42)	15.2–14.2	1.06	Reference group	0.43	137 (57)	17.9–18.6	−0.72	Reference group	0.07
>75	70 (34)	14.1–14.1	0.14	−1.05 (−2.67 to 0.56)		59 (25)	14.9–14.7	0.23	−0.84 (−0.67 to 2.37)	
Patient characteristic	Hip replacement (EQ-5D-3L index)					Knee replacement (EQ-5D-3L index)				
	n (%)	Mean Q1:QR	Mean difference	Adjusted difference	*p* value	*n* (%)	Mean Q1-QR	Mean difference	Adjusted difference	*p* value
SES (quintiles)				*Adjusted by age					*Adjusted by age	
1 (Least deprived)	19 (8.6)	0.17–0.29	−0.12	−0.08 (−0.25 to 0.09)		28 (11)	0.33–0.41	−0.07	−0.08(−0.22 to 0.07)	
2	40 (18.2)	0.22–0.17	0.05	0.08 (−0.05 to 0.21)		45 (17.7)	0.39–0.39	0	−0.02 (−0.15 to 0.10)	
3	62 (28.2)	0.22–0.25	−0.02	Reference group	0.39	50 (19.7)	0.32–0.29	0.02	Reference group	0.12
4	55(25)	0.28–0.28	−0.002	0.03 (−0.09 to 0.14)		71 (28.0)	0.31–0.33	−0.02	−0.05 (−0.16 to 0.06)	
5 (Most deprived)	44 (20)	0.17–0.14	0.03	0.04 (−0.08 to 0.17)		60 (23.6)	0.32–0.23	−0.09	0.06 (−0.06 to 0.18)	
Age (years)				*Adjusted by SES					*Adjusted by SES	
≤60	51 (23.2)	0.30–0.29	0.01	−0.02 (−0.13 to 0.09)		46 (18.1)	0.26–0.18	0.07	0.09 (−0.01 to 0.20)	
61–75	95 (43.2)	0.24–0.21	0.03	Reference group	0.23	147(57.9)	0.35–0.39	−0.04	Reference group	0.04
>75	74 (33.6)	0.14–0.20	−0.06	−0.08 (−0.18 to 0.01)		61 (24)	0.34–0.26	0.08	0.10 (0.01 to 0.20)	

OHS Oxford Hip Score, *OKS* Oxford Knee Score

for disease-specific PROMs was stronger than for generic PROMs. One explanation for this is that generic measures tend to have a more restricted range of responses, leading to greater homogeneity (smaller between-patient variability) in scores. ICCs define agreement between scores (within patients) in relative terms, so smaller population variation in scores will necessarily limit the strength of agreement.

These results suggest the main factors that may influence the differences between contemporary and retrospective reports, namely recall bias and response shift (a change in perception that can occur when circumstances change), did not have a significant influence. This may partly reflect the short time interval between measurements. Recall bias may arise when details of events go unnoticed and are not stored; new information is added to stored memories altering the details; and, over time, events are systematically distorted [26]. Such bias is influenced by the time between the event and its assessment: the longer the interval, the higher the probability of recall bias [27].

The lack of association between agreement and the length of the recall time in our results suggests that recall bias was minimal. It may be the case, as implicit theories of memory suggest, that the act of asking people to recall how they were before their surgery provided an anchor of their pre-surgical condition and hence formed the basis for stable recollection [28]. There is also a possibility that the exposure to a prior PROMs questionnaire could have aided recall. However, as an event in the patient's life, this is likely to pale in comparison with the subsequent hospital admission and operation in terms of a 'significant event' in the process of aiding the anchoring and assisting recollection of the patient's prior health.

The weaker agreement observed with the EQ-5D-3L is consistent with two previous studies that showed only moderate agreement when using PROMs with categorical data rather than continuous data [6, 7]. Lingard et al. [7] found this when items were not evenly distributed (i.e. when responses are clustered to at the severe end of the scales, e.g. severe pain and limited function).

As in this study, two previous studies observed the strength of agreement was high across age groups but decreased slightly with increasing age: OHS ICC for under 65 years 0.95 versus 0.85 for those older [8]; Western Ontario & McMaster Osteoarthritis Index for knee pain under 75 years 0.57 versus 0.47 for those older [7].

Strengths and limitations

This is the second largest such study ever undertaken, in addition to assessing agreement with ICCs which allowed differentiation between perfect agreement, systematic and random bias [29]. Bland–Altman plots [20] have provided a visual display of systematic bias or differences in relation to the scales of the PROMs providing an additional layer of understanding.

The one potential limitation concerns the representativeness of the sample who participated. Although they were broadly similar to the population of patients undergoing arthroplasty in England, they may have differed as regards some other unmeasured variables. It is possible that people who agreed to participate were more consistent in their recalled reports than the general population of patients.

Implications

These findings support the use of retrospective PROMs to obtain a baseline assessment of health status when contemporary collection is not feasible such as with emergency hospital admissions. In addition, retrospective collection offers an alternative even when contemporary is possible, an option that could not only facilitate higher participation rates but also lower the cost of data collection.

While this study has demonstrated the feasibility of collecting retrospective PROMs in patients who are recovering from an elective procedure (and who have already agreed to participate in a pre-operative contemporary report), research is now needed to determine the feasibility in emergency admissions. The latter have experienced an unexpected, sudden episode of illness and may still be unwell some days later. Whether collection of retrospective PROMs is feasible needs to be investigated.

Acknowledgements The authors are grateful to the staff and patients at the Barking, Havering and Redbridge University Hospitals NHS Trust (Site lead: Mr Mahbub Alam), Barts Health NHS Trust (Site lead: Mr Rej Bumbra, Jamila Kassam), Royal Free London NHS FT (Site lead: Mr Joyti Saksena, Ursula Knight) and the University College London Hospitals NHS FT (Site lead: Professor Fares Haddad, Ameena Hare).

Author contributions EK was the Principal Investigator of the study. NB is the Doctoral Supervisor. EK and JN performed the statistical analysis. EK and NB wrote the paper with input from JN.

Funding The research and EK is funded by Economic & Social Research Council doctoral fellowship. Grant Reference: ES/J500021/1. The research was supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care North Thames at Barts Health NHS Trust. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

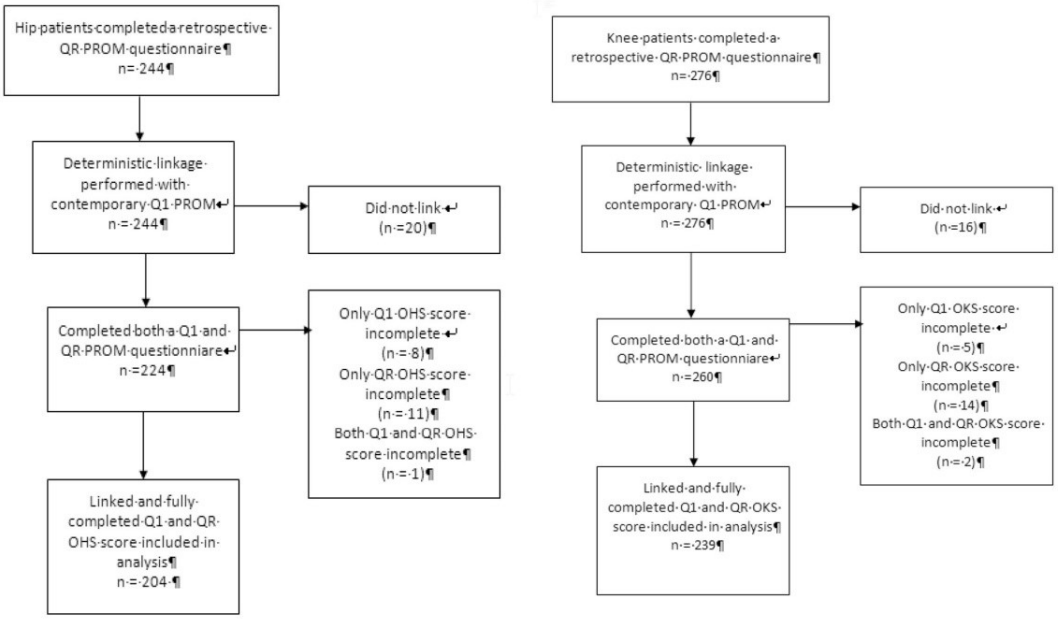
Compliance with ethical standards

Conflict of interest We declare that we have no conflicts of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and national research committee (NHS Health Research Authority) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

Appendix: study flow



References

1. Black, N. (2013). Patient reported outcome measures could help transform healthcare. *BMJ: British Medical Journal, 346*, 19–21.
2. Greenhalgh, J., Dalkin, S., Gooding, K., Gibbons, E., Wright, J., Meads, D. et al. (2017). Functionality and feedback: A realist synthesis of the collation, interpretation and utilisation of patient-reported outcome measures data to improve patient care. *Health Services and Delivery Research, 5*(2), 1–163.
3. Varagunam, M., Hutchings, A., Neuburger, J., & Black, N. (2014). Impact on hospital performance of introducing routine patient reported outcome measures (PROMs) in surgery. *Journal of Health Services Research & Policy, 19*, 77–84.
4. NHS England. (2016). A&E Attendances and Emergency Admissions. May 2016 Monthly Report. https://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2016/06/Monthly-AE-Report-May-16.pdf.
5. Kwong, E., & Black, N. (2017). Retrospectively patient-reported pre-event health status showed strong association and agreement with contemporaneous reports. *Journal of Clinical Epidemiology, 81*, 22–32.
6. Emberton, M., Challands, A., Styles, R. A., Wightman, J. A. K., & Black, N. (1995). Recollected versus contemporary patient reports of pre-operative symptoms in men undergoing transurethral prostatic resection for benign disease. *Journal of Clinical Epidemiology, 48*(6), 749–756.
7. Lingard, E. A., Wright, E. A., & Sledge, C. B. (2001). Pitfalls of using patient recall to derive preoperative status in outcome studies of total knee arthroplasty. *The Journal of Bone and Joint Surgery. American Volume, 83*(8), 1149–1156.
8. Howell, J., Xu, M., Duncan, C. P., Masri, B. A., & Garbuz, D. S. (2008). A comparison between patient recall and concurrent measurement of preoperative quality of life outcome in total hip arthroplasty. *Journal of Arthroplasty, 23*(6), 843–849.
9. Communities and Local Government. (2007). *The English Indices of Deprivation 2007* (Vol. 59). Communities and Local Government Publications.
10. Dawson, J., Fitzpatrick, R., & Carr, A. (1996). Questionnaire on the perceptions of patients about total hip replacement. *Journal of Bone and Joint Surgery, 78*, 185–190.
11. Dawson, J., Fitzpatrick, R., & Murray, D. (1998). Questionnaire on the perceptions of patients about total knee replacement. *Journal of Bone and Joint Surgery, 80*, 63–69.
12. Pynsent, P. B., Adams, D. J., & Disney, S. P. (2005). The Oxford hip and knee outcome questionnaires for arthroplasty. *The Journal of Bone and Joint Surgery. British Volume, 87 SRC-*, 241–248.
13. EuroQol Group. (1990). EuroQol—A new facility for the measurement of health-related quality of life. *Health Policy, 16*, 199–208.
14. Dolan, P., & Roberts, J. (2002). Modelling valuations for Eq-5d health states: An alternative model using differences in valuations. *Medical Care, 40*, 442–446.
15. Feng, Y., Parkin, D., & Devlin, N. J. (2014). Assessing the performance of the EQ-VAS in the NHS PROMs programme. *Quality of Life Research, 23*, 977–989.
16. McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlations coefficients. *Psychological Methods, 1*(1), 30–46.
17. Streiner, D., & Norman, G. (2008). *Health measurement scales: Guide to their development and use* (4th ed.). Oxford: Oxford University Press.
18. Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159–174.
19. Mander, A. (2012, June 17). BATPLOT: Stata module to produce Bland-Altman plots accounting for trend. Retrieved from https://econpapers.repec.org/RePEc:boc:bocode:s448703.
20. Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet, 1*(8476), 307–310.
21. Neuburger, J., Hutchings, A., Allwood, D., Black, N., & Van der Meulen, J. (2012). Sociodemographic differences in the severity and duration of disease in patients undergoing hip or knee replacement surgery. *Journal of Public Health, 34*, 421–429.
22. Hutchings, A., Neuburger, J., Gross Frie, K., Van der Meulen, J., & Black, N. (2012). Factors associated with non-response in routine use of patient reported outcome measures after elective surgery in England. *Health and Quality of Life Outcomes, 10*, 34.
23. Bryant, D., Norman, G., Stratford, P., Marx, R. G., Walter, S. D., & Guyatt, G. (2006). Patients undergoing knee surgery provided accurate ratings of preoperative quality of life and function 2 weeks after surgery. *Journal of Clinical Epidemiology, 59*(9), 984–993.
24. Marsh, J., Bryant, D., & MacDonald, S. J. (2009). Older patients can accurately recall their preoperative health status six weeks following total hip arthroplasty. *Journal of Bone and Joint, 91*, 2827–2837.
25. Helfand, B. T., Fought, A., Manvar, A. M., & McVary, K. T. (2010). Determining the utility of recalled lower urinary tract symptoms. *Urology, 76*(2), 442–447.
26. Schmier, J. K., & Halpern, M. T. (2004). Patient recall and recall bias of health state and health status. *Expert Review of Pharmacoeconomics & Outcomes Research, 4*(2), 159–163.
27. Skowronski, J. J., Betz, A. L., & Thompson, C. P. (1991). Social memory in everyday life: Recall of self-events and other-events. *Journal of Personality and Social Psychology, 64*, 831–843.
28. Baddeley, A. (1997). *Human memory. Theory and practice* (Revised ed.). Hove: Psychology Press.
29. Rousson, V., Gasser, T., & Seifert, B. (2002). Assessing intrarater, interrater and test-retest reliability of continuous measurements. *Statistics in Medicine, 21*(22), 3431–3446.

Chapter 4: Using GPPS survey data as a source of baseline PROM scores: Methods for matching with patient cohorts.

This chapter presents analyses exploring the potential for using population-based PROMs (the EQ-5D-3L Index Score) from the GP Patient Survey (GPPS) to form baseline PROM scores for patient cohorts. It explores alternative methods of matching GPPS populations to elective patient cohorts, and different ways of comparing GPPS values to retrospective and contemporary PROMs. For this chapter, I have compared surgical patients from the elective study cohort (Chapter 3) to the GPPS population. The motivation for this study was to evaluate whether EQ-5D scores from the GPPS population sample could be harnessed for use to form a proxy patient baseline PROM in the place of a retrospective PROM, potentially decreasing the administrative burden of collecting PROMs in emergency care settings. I developed the methods in the study design with support from Professor Nick Black and Dr Gary Abel (University of Exeter); the analysis was conducted with the statistical advice from Dr Jenny Neuburger and Dr Gary Abel. Professor Nick Black provided feedback and comments on the layout of this chapter. This has not been prepared for submission to a journal.

4.1 Introduction

An alternative to using recall data to obtain baseline health status for patients admitted for an emergency condition (i.e. before the sudden and unexpected onset of ill health) could be to use PROMS in groups with similar characteristics drawn from population surveys such as the GP Patient Survey (GPPS). Matched data on patient reported health status could be used as a form of proxy measure of baseline health status in emergency patient cohorts. If feasible, it could offer a cheaper alternative compared with collecting retrospective PROMs.

4.1.1 What we already know

Previous studies have compared retrospectively collected PROMs with measures from general population surveys [1]. All of these studies used general population values, but none attempted to match the population sample to that of the patients with the exception of matching on age.

The GP Patient Survey (GPPS) is an annual survey with a questionnaire on patient experience mailed every year to approximately 2.2 million adults who are registered with a GP in England, with a response rate of just under 40% providing a sample of around 800,000. Since 2011, the EQ-5D-3L (population health status) has been included [2].

It is known that reported EQ-5D scores of the general population may be very different from those of elective surgical patients. The EQ-5D index population value for England for someone aged between 65-74 is 0.73 [3] compared with 0.36 for elective hip arthroplasty patients with a median age of 67 [4,5]. However, such comparisons are typically only adjusted for age. The GPPS collects data on several other patient characteristics including postal address, from which area-based socioeconomic status (SES) can be derived, sex, and self-reported co-morbidities. Therefore this provides the opportunity to conduct more detailed matching than just using the overall population value. It may be possible that with the matching of patient characteristics, differences between the population's health status and that of surgical patients would be less. If so, population health status could be used in place of retrospective questionnaires to obtain a baseline score, which could have both clinical and cost advantages,

including reducing patient and staff burden of collecting baseline PROM questionnaires.

My aim was to investigate whether mean general population EQ-5D scores from the GPPS could be used in place of contemporaneously or retrospectively patient reported baseline EQ-5D scores using matching by patient characteristics. Furthermore, I explored whether different levels of specificity used in the matching process enables GPPS to be more similar to baseline retrospective and contemporary PROMs. If such methods are feasible, I aimed to see whether this provides a way to harness available GPPS EQ-5D data to use in comparisons to support the auditing of health services.

Four objectives were identified to achieve these aims:

1) To compare the contemporary and retrospective self-reported health status (mean EQ-5D scores) of a cohort of elective surgical patients with that of the general population of England matched for sex, age, SES and number of comorbidities.
2) To test whether additional matching by geographical location reduces the differences between population and self-reported mean scores.
3) To test whether additional matching for specific co-morbidities reduces the differences between mean scores.
4) To test whether different ways of handling patients' primary condition reduces the differences between mean scores.

4.2 Methods

Elective surgical patients who participated in the study reported in Chapter 3 were matched to GPPS data (held at the University of Exeter) using patient characteristics available in both data sets. One-to-many matching was conducted with one surgical patient matched to as many GPPS respondents as fitted the matching criteria. The mean GPPS EQ-5D score per match was used as the comparison between patients' reported contemporary (Q1) and retrospective (QR) EQ-5D and GPPS EQ-5D.

4.2.1 Population sample from GPPS

Data from the 2011–2012 GPPS included the EQ-5D-3L, the same version as that used for the patient cohort. 2.7 million patients were surveyed, with a 38% response rate resulting in a sample of approximately 1 million. Patients in the GPPS are randomly sampled, with stratification, from the contact records of all general practices across England. Questionnaires were sent in July 2011 and January 2012 to approximately 1.40 and 1.36 million patients, respectively. Non-responders were mailed reminders after two months following the initial questionnaire [6].

Alongside patient experience items, patients were asked to report any long-standing health condition from a list of twelve common conditions: angina or long-term heart problem, arthritis or long-term joint problem, asthma or long-term chest problem, cancer in the last 5 years, deafness or severe hearing impairment, diabetes, epilepsy, high blood pressure, kidney or liver disease, long-term back problem, long-term mental health problem, long-term neurological problem and 'another' long-term condition [7].

4.2.2 EQ-5D in patient cohort and in GPPS

The EQ-5D-3L version is a generic utility score of health state derived from individuals' responses, on a three-level ordinal scale (no problems, moderate problems and severe problems), for each of its five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) [8]. UK tariffs were used to obtain an index score which ranges from −0.59 (the worst possible health state) to 1 (indicates best possible health state). The value of 0 is equal to death and negative values represent health states worse than death [9].

4.2.3 Patient cohort

As described in more detail in Chapter 3, study participants were patients undergoing hip arthroplasty (primary operation or revision surgery) from four hospitals. Their baseline contemporary (Q1) EQ-5D-3L mean scores were similar to that for all patients' included in the National PROMs Programme in England [10]. The contemporary PROM questionnaire (Q1) contained

questions about the presence of any long-standing health conditions; patients are asked to self-report these from a list of twelve common conditions: heart disease (for example angina, heart attack or heart failure), high blood pressure, problems caused by a stroke, leg pain due to poor circulation, lung disease, diabetes, kidney disease, liver disease, cancer (within in the last 5 years), diseases of the nervous system (for example Parkinson's disease or multiple sclerosis), depression.

Patients also completed a retrospective PROM questionnaire (QR) in the immediate post-operative period prior to discharge.

4.2.4 Matching patient cohort to population sample

The patient cohort was matched to GPPS population on the following characteristics: sex, age, socioeconomic status, and co-morbidities. The sample sizes were large enough, relative to the number of matching characteristics, to permit exact one-to-many matching. The following variables were created to carry out the matching. Patients' ages were grouped into the following eight categories: 18-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75-84, and 85+ years. Socioeconomic status (SES) was derived from the Index of Multiple Deprivation (IMD) of a patient's local area (LSOAs) based on postcode, which was then grouped into quintiles based on the national ranking of areas by IMD to match the GPPS variable.

In a second exercise, matching was restricted by geographical location, with patients matched only to GPPS respondents living in the same local authority, to test whether this reduced differences between population and patient cohorts' mean EQ-5D scores.

Finally, I investigated whether different ways of handling co-morbidities data altered any difference in mean EQ-5D score between the patient cohort and the population. For co-morbidities, conditions in the patient cohort were mapped onto the population (Table 4-1). Two ways were used for matching co-morbidities; first by a simple count of the number of comorbid conditions and second by exact matching to explore whether greater specificity narrowed any differences between population and patient EQ-5D mean scores.

Different ways of handling patients' most likely primary condition (i.e. arthritis) within self-reported comorbidities were explored using three different methods. The first method takes the data at face value: those with arthritis reported by patients are matched with GPPS respondents with arthritis while those patients not reporting arthritis are matched with GPPS respondents also not reporting arthritis. The second method treats the entire patient cohort as having arthritis as their primary condition, whether or not this was self-reported on the questionnaire. Patients were then matched to only those in the GPPS population reporting arthritis or a long-term joint problem. The third method disregarded arthritis in both data sets.

After matching, descriptive analyses were conducted by stratifying patient characteristics to observe the differences in mean EQ-5D scores for the different matching methods. Data were stratified by patient characteristics and z tests carried out to compare differences in mean EQ-5D scores between GPPS, and contemporary PROM and retrospective PROM using the matching methods described above.

4.3 Results

There were 203 elective patients available for matching, of whom 21 could not be matched: 20 had missing data on co-morbidities and one had no postcode. The mean EQ-5D score for the patient cohort was compared with mean EQ-5D for the population samples using three different matching strategies; age, sex, SES and number of comorbidities; age, sex, SES, number of comorbidities and local authority; and age, sex, SES and specific comorbidities. In addition, different ways of handling the patients' primary condition (arthritis) were explored, as described in the methods above.

4.3.1 Contemporary and retrospective EQ-5D scores of patients matched for age, sex, SES and number of comorbidities

When matched for age, sex, SES, and number of co-morbidities, the national population mean EQ-5D was 0.68 (SD 0.27), whereas the elective patient cohort's mean Q1 was 0.24 (SD 0.33) and mean QR was 0.22 (SD 0.35). The

differences in means between population and patients were 0.44 and 0.46 respectively.

Differences in mean EQ-5D scores between matched patient and GPPS groups were not much smaller, although differences varied by matching characteristic (age, sex and SES). Comparisons by age and sex (Table 4-2), the differences were 0.40-0.53 between Q1 and GPPS and 0.38-0.56 between QR and GPPS. Table 4-3 shows comparisons presented by SES groups for men and women separately. The differences were between 0.45-0.58 and 0.38-0.60 respectively for each group.

4.3.2 Contemporary and retrospective EQ-5D scores of patients matched for age, sex, SES, number of comorbidities and local authority

These differences were not reduced when matching was restricted to patients and GPPS respondents living in the same local authority: differences between the GPPS mean EQ-5D score ranged from 0.38-0.56 for Q1 and 0.38-0.55 for QR. Comparisons in the differences between national and local authority matched means by age and sex are shown in Table 4-4. Results shown in Table 4-5 compares the differences between national and local authority matched means presented by sex and SES groups. Again these differences were not narrowed when matching restriction by local authority was used: 0.34-0.57 for Q1 and 0.41-0.60 for QR. Although sample sizes were smaller when restricted by local authority, the differences between the mean scores remained consistently statistically significant.

4.3.3 Contemporary and retrospective self-reported EQ-5D scores of patients matched for age, sex, SES and specific comorbidities

Matching by specific comorbidities did not change the extent of the differences between population and patient cohort EQ-5D scores compared to when a simple count of comorbidities was used (Table 4-6). For Q1 the differences were 0.40-0.53 for the total count and 0.46-0.59 for specific comorbidities. For QR the differences were 0.38-0.56 using the total count and 0.45-0.59 using specific comorbidities.

4.3.4 Different ways of handling patients' primary condition

Table 4-7 shows the differences (by sex and SES groups) using different methods for handling the primary condition for the patient cohort (arthritis). The difference between GPPS and patients' mean EQ-5D is similar when patients reporting arthritis are matched with GPPS respondents with arthritis (Method 1: differences in means were 0.40-0.53 for Q1 and 0.38-0.56 for QR), compared to when arthritis was disregarded in both data sets (Method 3: differences in means were 0.44-0.56 for Q1 and 0.42-0.58 for QR).

There was a slight decrease in the differences between population and patient cohort scores when all study participants were treated as having arthritis as their primary condition regardless whether this was reported in their co-morbidities (Method 2). The differences in means were 0.21-0.42 for Q1 and 0.26-0.41 for QR. However, these differences remained statistically different.

4.4 Discussion

4.4.1 Main Findings

Patients' EQ-5D mean baseline score was very different from that of the general population who responded to the GPPS. The elective patient cohort PROMs scores were consistently lower than the GPPS mean scores. This remained so even after using several patient characteristics to match the population sample including age, sex, SES and number of co-morbidities. Matching by local authority did not reduce the differences in mean scores, compared to using the national dataset. Therefore, the use of the latter would be preferable as a larger sample is achievable. More specific matching of co-morbidities to exact conditions also did not narrow the differences in mean EQ-5D than simply using a count of the number of co-morbidities. The latter again has the advantage of providing a larger sample for one-to-many matching.

Making the assumption that all study patients had arthritis when considering data on co-morbidities provided smaller differences between patients and the population sample compared to the other methods of handling comorbidities. However, this did not narrow the differences between population values and the patient cohort to a statistically significant level. Whether this observation

holds true for other conditions leading to emergency admissions needs to be established. In certain emergency conditions, such as AMI, there is both a clear underlying primary condition and a corresponding recorded co-morbidity on GPPS, making it possible to investigate this further. However, this may not be possible in instances when patients with different conditions undergo the same emergency operation and not all the primary conditions are captured in the GPPS data.

4.4.2 Strengths and Limitations of GPPS data

The national GPPS is one of the largest annual surveys of patients in the world. It provides an overview of the experience and quality of care provided by general practices in England [11]. It was carefully developed, with expert and stakeholder advice, and piloted prior to its routine use in England [6].

For GPPS data from 2011-2012, a total of 1,037,946 people (38%) returned questionnaires, comparable to that achieved in other surveys using a similar methodology in the UK [6]. Co-morbidity data were reported by 906,578 (87.3%), EQ-5D scores were complete for 831,537 respondents (80.1%) and only 574 (<0.001%) had missing information on deprivation [12].

Certain socio-demographic factors predict GPPS response. Younger patients (age 18-29 years) were the least likely to respond. However, for the patient groups that are similar to hip arthroplasty patients (middle age to elderly), responders were representative of the general population. This is because response rates increase substantially to a peak at the ages 70-79, where the odds of responding were 5.5 times as high as for those aged 18-29. For socioeconomic deprivation, the odds of responding declined approximately linearly with increasing deprivation. The odds of responding were 41% lower for men than for women, after allowing for the effects of age and deprivation [11].

Limitations could arise from non-participation and item non-response. Women, the middle-aged and those in affluent areas are more likely to respond. Non-response may pose an issue when using GPPS data as a population norm if non-respondents differ from respondents in their health status. Although

certain groups are under-represented (younger, men, ethnic minorities, socially deprived), with the exception of ethnicity, these were characteristics that were matched with the patient cohort in this exploratory investigation. One minor limitation was that age was recorded in ten-year bands for GPPS, necessitating similar groupings for the patient cohort.

Comorbidity data may be limited by whether patients' self-reported conditions correspond with objective health status measures. However, the prevalence of individual co-morbidities reported in GPPS was similar to comparable reports from other sources except for diabetes, high blood pressure and long-term back problems. It has been suggested that other possible explanations for these differences may be due to comparative data being out of date [12].

For any non-response bias to affect EQ-5D scores, the association between health status and the likelihood of responding to GPPS would need to be associated with unobserved and thus unmatched characteristics such as ethnic group or educational attainment [13]. Although, it is not possible to estimate the impact of bias due to these variables, published meta-analyses on probability sampled surveys suggest that response rates are not a strong predictor for any non-response bias [14].

4.4.3 Implications for further research

These exploratory findings provide an insight into the potential of using matching by patient characteristics, made possible because these characteristics have been routinely collected within the annual GPPS. This approach has been shown not to be suitable for elective patients, even after matching, as their baseline health status remains significantly worse than that of populations with similar characteristics. This is most likely due to the presence of long-standing medical conditions that are severe enough to warrant major surgery.

The analyses presented in this chapter are not supportive of the use of GPPS PROMs for forming baseline PROMs for patient cohorts. However, there are a number of reasons to think that this approach may nevertheless be of value in some cohorts of emergency patients, especially those who have a sudden

illness or condition without prior long-standing ill health. In these patients, their baseline health status may be closer to those of the general population and hence matching to GPPS data could provide a suitable alternative to collecting retrospective PROMs. It would therefore be useful to explore this option in emergency patient cohorts.

Specifically, with AMI, whether handling the primary condition by treating all patients as having an underlying heart condition and matching to those in GPPS with heart disease when considering data on co-morbidities, should be investigated.

Although direct comparisons between GPPS data with contemporary PROMs is not possible in cohorts of emergency admissions, it would be useful to compare retrospective baseline PROMs of emergency admission patients with GPPS population norms using the matching methods explored in this chapter.

4.5 Tables

Table 4-1 Co-morbidities reported in the surgical questionnaire and the GPPS

Study Condition	GPPS Condition
Heart disease	Angina or long-term heart problem
Arthritis	Arthritis or long-term joint problem
Lung disease	Asthma or long-term chest problem
Cancer	Cancer in the last 5 years
Diabetes	Diabetes
High blood pressure	High blood pressure
Kidney or Liver disease	Kidney or liver disease
Depression	Long-term mental health problem
Nervous system	Long-term neurological problem

Table 4-2 Differences between contemporary (Q1) & retrospective (QR) patients' mean EQ-5D and national GPPS mean EQ-5D overall and by age and sex

Patients by Sex & Age groups	Differences in means GPPS vs. Q1 (95% CI)	Differences in means GPPS vs. QR (95% CI)
Overall	0.44 (0.39-0.49)	0.46 (0.40-0.51)
Men, 60 or under	0.40 (0.23 - 0.56)	0.48 (0.33 - 0.63)
Men, 61-75	0.47 (0.33 - 0.61)	0.53 (0.40 - 0.67)
Men, 76 and above	0.53 (0.40 - 0.67)	0.38 (0.22 - 0.53)
Women, 60 or under	0.48 (0.34 - 0.62)	0.51 (0.38 - 0.65)
Women, 61-75	0.52 (0.43 - 0.59)	0.56 (0.48 - 0.64)
Women, 76 and above	0.50 (0.41 -0.60)	0.48 (0.40 - 0.58)

* 95% CI calculated using diff +/- 1.96*SE(diff). SE(diff) = sqrt(SD_q1^2/n1 + SD_q2^2/n2)

Table 4-3 Differences between contemporary (Q1) & retrospective (QR) patients' mean EQ-5D and national GPPS mean EQ-5D by sex and SES

Patients by Sex & SES (quintiles) groups	Differences in means GPPS vs Q1 (95% CI)	Differences in means GPPS vs QR (95% CI)
Women, 1 (least deprived)	0.45 (0.13-0.77)	0.60 (0.40-0.81)
Women, 2	0.52 (0.38-0.67)	0.58 (0.44-0.72)
Women, 3	0.58 (0.46 - 0.70)	0.56 (0.45 - 0.67)
Women, 4	0.46 (0.37 - 0.56)	0.47 (0.37 - 0.58)
Women, 5 (most deprived)	0.53 (0.41 - 0.64)	0.55 (0.43 - 0.68)
Men, 1 (least deprived)	0.47 (0.48 - 0.91)	0.47 (0.18 - 0.77)
Men, 2	0.49 (0.18 - 0.81)	0.49 (0.18 - 0.80)
Men, 3	0.45 (0.31- 0.60)	0.47 (0.30 - 0.63)
Men, 4	0.32 (0.15 - 0.51)	0.39 (0.24 - 0.55)
Men, 5 (most deprived)	0.48 (0.29 - 0.68)	0.56 (0.40 - 0.73)

Table 4-4 Differences between national and local authority GPPS and patients' mean EQ-5D by age and sex

	Difference in Means GPPS vs Q1		Difference in Means GPPS vs QR	
Patients by Sex & Age	**National** (95% CI)	**Local authority** (95% CI)	**National** (95% CI)	**Local authority** (95% CI)
Men, 60 or under	0.40 (0.23 - 0.56)	0.48 (0.33 - 0.63)	0.40 (0.23 - 0.56)	0.48 (0.33 - 0.63)
Men, 61-75	0.47 (0.33 - 0.61)	0.53 (0.40 - 0.67)	0.48 (0.34 - 0.62)	0.55 (0.41 - 0.68)
Men, 76 and above	0.53 (0.40 - 0.67)	0.38 (0.22 - 0.53)	0.53 (0.40 - 0.67)	0.38 (0.32 - 0.63)
Women, 60 or under	0.48 (0.34 - 0.62)	0.51 (0.38 - 0.65)	0.46 (0.35 - 0.58)	0.51 (0.37 - 0.64)
Women, 61-75	0.52 (0.43 - 0.59)	0.56 (0.48 - 0.64)	0.51 (0.43 - 0.59)	0.55 (0.47 - 0.64)
Women, 76 and above	0.50 (0.41 - 0.60)	0.48 (0.40 - 0.58)	0.50 (0.40 - 0.59)	0.48 (0.39 - 0.57)

Table 4-5 Differences between national and local authority GPPS and patients' mean EQ-5D by age and sex

	Difference in Means GPPS vs Q1		Difference in Means GPPS vs QR	
Patients by Sex & SES (quintiles) groups	**National (95% CI)**	**Local authority (95% CI)**	**National (95% CI)**	**Local authority (95% CI)**
Women, 1 (least deprived)	0.45 (0.13 - 0.77)	0.44 (0.13 - 0.77)	0.60 (0.40 - 0.81)	0.60 (0.40 - 0.81)
Women, 2	0.52 (0.38 - 0.67)	0.51 (0.37 - 0.66)	0.58 (0.44 - 0.72)	0.57 (0.43 - 0.71)
Women, 3	0.58 (0.46 - 0.70)	0.57 (0.46 - 0.68)	0.56 (0.45 - 0.67)	0.55 (0.44 - 0.66)
Women, 4	0.46 (0.37 - 0.56)	0.46 (0.38 - 0.56)	0.47 (0.37 - 0.58)	0.47 (0.37 - 0.58)
Women, 5 (most deprived)	0.53 (0.41 - 0.64)	0.55 (0.44 - 0.66)	0.55 (0.43 - 0.68)	0.57 (0.45 - 0.70)
Men, 1 (least deprived)	0.47 (0.48 - 0.91)	0.47 (0.48 - 0.91)	0.47 (0.18 - 0.77)	0.47 (0.18 - 0.77)
Men, 2	0.49 (0.18 - 0.81)	0.49 (0.18 - 0.81)	0.49 (0.18 - 0.80)	0.49 (0.18 - 0.80)
Men, 3	0.45 (0 .31 - 0.60)	0.47 (0.32 - 0.61)	0.47 (0.30 - 0.63)	0.48 (0.32 - 0.64)
Men, 4	0.32 (0.15 - 0.51)	0.34 (0.19 - 0.50)	0.39 (0.24 - 0.55)	0.41 (0.26 - 0.57)
Men, 5 (most deprived)	0.48 (0.29 - 0.68)	0.52 (0.34 - 0.71)	0.56 (0.40 - 0.73)	0.60 (0.43 - 0.76)

Table 4-6 Differences between national GPPS and patients' mean EQ-5D by age and sex groups using comorbidity matched by number and by specific condition

	Difference in Means GPPS vs Q1		Difference in Means GPPS vs QR	
Patients by Sex & Age	**Co-morbidities by number**	**Co-morbidities by specific matching**	**Co-morbidities by number**	**Co-morbidities by specific matching**
Men, 60 or under	0.40 (0.23 - 0.56)	0.46 (0.31 - 0.61)	0.48 (0.33 - 0.63)	0.53 (0.38 - 0.68)
Men, 61-75	0.47 (0.33 - 0.61)	0.50 (0.37 - 0.64)	0.53 (0.40 - 0.67)	0.59 (0.45 - 0.72)
Men, 76 and above	0.53 (0.40 - 0.67)	0.59 (0.46 - 0.72)	0.38 (0.22 - 0.53)	0.45(0.29 - 0.60)
Women, 60 or under	0.48 (0.34 - 0.62)	0.53 (0.28 - 0.80)	0.51 (0.38 - 0.65)	0.56 (0.43 - 0.69)
Women, 61-75	0.52 (0.43 - 0.59)	0.51 (0.44 - 0.60)	0.56 (0.48 - 0.64)	0.54 (0.47 - 0.63)
Women, 76 and above	0.50 (0.41 - 0.60)	0.50 (0.41 - 0.60)	0.48 (0.40 - 0.58)	0.48 (0.40 - 0.57)

Table 4-7 Differences between national GPPS and patients' mean EQ-5D by age and sex groups using different methods of handling the primary condition

	Difference in Means GPPS vs Q1 (95% CI)			Difference in Means GPPS vs QR (95% CI)		
Patients by Sex & Age	**Method 1**	**Method 2**	**Method 3**	**Method 1**	**Method 2**	**Method 3**
Men, 60 or under	0.40 (0.23 - 0.56)	0.21 (0.05 - 0.38)	0.45 (0.28 - 0.62)	0.48 (0.33 - 0.63)	0.26 (0.11 - 0.41)	0.52 (0.37 - 0.67)
Men, 61-75	0.47 (0.33 - 0.61)	0.29 (0.16 - 0.43)	0.48 (0.33 - 0.62)	0.53 (0.40 - 0.67)	0.35 (0.22 - 0.48)	0.56 (0.43 - 0.70)
Men, 76 and above	0.53 (0.40 - 0.67)	0.41 (0.29 - 0.54)	0.44 (0.32 - 0.58)	0.38 (0.22 - 0.53)	0.26 0.11 - 0.41)	0.42 (0.26 - 0.57)
Women, 60 or under	0.48 (0.34 - 0.62)	0.31 (0.17 - 0.44)	0.56 (0.43 - 0.70)	0.51 (0.38 - 0.65)	0.32 (0.19 - 0.45)	0.58 (0.44 - 0.71)
Women, 61-75	0.52 (0.43 - 0.59)	0.40 (0.32 - 0.47)	0.55 (0.42 - 0.67)	0.56 (0.48 - 0.64)	0.41 (0.33 - 0.49)	0.58 (0.50 - 0.66)
Women, 76 and above	0.50 (0.41 - 0.60)	0.42 (0.33 - 0.51)	0.54 (0.45 - 0.63)	0.48 (0.40 - 0.58)	0.39 (0.30 - 0.48)	0.58 (0.43 - 0.61)

4.6 References

1 Kwong E, Black N, J. Retrospectively collected patient reported outcomes show strong association and agreement with their contemporaneous reports. *J Clin Epidemiol* 2017;**81**:22–32.

2 NHS England. GP Patient Survey. 2017. Available at: https://www.england.nhs.uk/statistic.

3 Janssen B, Szende A. Self-Reported Population Health: An International Perspective based on EQ-5D. Springer Netherlands 2014. 19–30.

4 Neuburger J, Hutchings A, Allwood D, *et al.* Sociodemographic differences in the severity and duration of disease in patients undergoing hip or knee replacement surgery. *J Public Heal* 2012;**34**:421–9.

5 Varagunam M, Hutchings A, Neuburger J, *et al.* Impact on hospital performance of introducing routine patient reported outcome measures (PROMs) in surgery. *J Health Serv Res Policy* 2014;**19**:77–84.

6 Campbell J, Smith P, Nissen S, *et al.* The GP Patient Survey for use in primary care in the National Health Service in the UK -- development and psychometric characteristics. *BMC Fam Pract* 2009;**10**:57.

7 NHS England. GPPS Surveys and Reports. 2017.www.gpatient.co.uk/download/Questionnaires/

8 EuroQol Group. EuroQol - a new facility for the measurement of health-related quality of life. *Health Policy (New York)* 1990;**16**:199–208.

9 Dolan P, Roberts J. Modelling valuations for Eq-5d health states: an alternative model using differences in valuations. *Med Care* 2002;**40**:442–6.

10 Kwong E, Neuburger J, Black N. Agreement between retrospectively and contemporaneously collected patient-reported outcome measures (PROMs) in hip and knee replacement patients. *Qual Life Res* Published Online First: February 2018.

11 Roland M, Elliott M, Lyratzopoulos G, *et al.* Reliability of patient

responses in pay for performance schemes: analysis of national General Practitioner Patient Survey data in England. BMJ. 2009;**339**.

12 Mujica-Mota RE, Roberts M, Abel G, *et al.* Common patterns of morbidity and multi-morbidity and their impact on health-related quality of life: evidence from a national survey. *Qual Life Res* 2015;**24**:909–18.

13 Burt J, Lloyd C, Campbell J, *et al.* Variations in GP–patient communication by ethnicity, age, and gender: evidence from a national primary care patient survey. Br. J. Gen. Pract. 2016;**66**:e47-52.

14 Groves RM, Peytcheva E. The Impact of Nonresponse Rates on Nonresponse BiasA Meta-Analysis. *Public Opin Q* 2008;**72**:167–89.

Chapter 5

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Esther Kwong
Principal Supervisor	Nick Black
Thesis Title	The use of patient reported outcome measures (PROMs) for evaluating emergency admissions

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

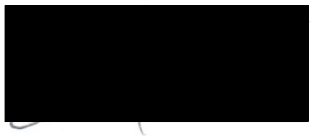
**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	Journal of Patient Reported Outcomes
Please list the paper's authors in the intended authorship order:	Esther Kwong, Nick Black
Stage of publication	**Undergoing revision**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	Please refer to details in the following page



Student Signature: _____ **Date:** 24 August 2018

Supervisor Signature: _____ **Date:** 24 August 2018

Chapter 5

Feasibility of collecting retrospective patient reported outcome measures (PROMs) in emergency hospital admissions

This chapter described the feasibility of recruiting patients and the collection of retrospective baseline PROMs with emergency admissions patients in NHS hospitals. The purpose was to test the acceptability of such in two diverse emergency patient groups, in a variety of hospital settings, to explore the relative merits and variables influenced by these factors. This was a cohort study in acute myocardial infarction patients (STEMI) and emergency laparotomy (EL) patients.

I lead the study design with the guidance of Professor Nick Black, in partnership with clinical collaborators from the respective conditions. For emergency laparotomy, a PROMs project group was established after meetings with National Clinical Emergency Laparotomy Audit chairs Professor Mike Grocott and Dr Dave Murray. The project group subsequently provided feedback about the study design. For STEMI, Professor Steffen Petersen and colleagues from the Cardiology department at Barts Health peer reviewed and provided feedback on the study design, protocol and IRAS application.

Unlike orthopaedic departments, collection of PROMs was novel for these areas, to ensure successful running of the studies, I held regular meetings with clinical teams at the sites prior to the start of data collection to adapt PROMs collection protocol to local processes. I designed and provided training in the form of written materials and recorded training videos to introduce the study, the consent and invitation processes for emergency patients with the support of a colleague (Ms Ursula Knight) from the orthopaedic studies (Chapter 3). I held a site confirmation meeting or telephone interview with site leads and study teams prior to the start of data collection at each site and provided on-going support, advice and quality assurance to data collection through regular meetings and contact with all the sites throughout the study period.

I prepared the first draft of this manuscript and made revisions following comments from my supervisor Nick Black. This paper has been submitted to the Journal of Patient Reported Outcomes.

Feasibility of collecting retrospective patient reported outcome measures (PROMs) in emergency hospital admissions

Esther Kwong

Nick Black

Authors Affiliation

Corresponding Author:

Dr Esther Kwong, MBBS, MSc, BSc, Doctoral Candidate,

Dept. of Health Services Research & Policy,

London School of Hygiene and Tropical Medicine

15-17 Tavistock Place, London WC1H 9SH, UK.

Email: esther.kwong@lshtm.ac.uk Phone: +44 (0)207 958 8285

Professor Sir Nick Black, Professor, Dept. of Health Services Research and Policy, London School of Hygiene and Tropical Medicine

Keywords

Patient Reported Outcome Measures; Health Status; Health-related quality of life; Retrospective; Feasibility, Emergency admissions, STEMI, Emergency Laparotomy

Contribution to the study

EK was the Principal Investigator of the study. NB is the Doctoral Supervisor. EK and NB wrote the paper with input from JN.

Conflicts of Interest

We declare that we have no conflicts of interest.

Funding

The research and EK is funded by Economic & Social Research Council doctoral fellowship. Grant Reference: ES/J500021/1

The research was supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care North Thames at Barts Health NHS Trust. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health."

Ethical approval and Informed Consent

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and national research committee (NHS Health Research Authority) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. NHS ethical approval obtained from South East Coast - Brighton & Sussex Research Ethics Committee (REC reference: 16/LO/2053). Informed consent was obtained from all individual participants included in the study.

Acknowledgements

We thank: The patients and staff of 16 hospitals that participated in the study. NELA project team. Site and Study leads.

Collaborators: Mike Grocott, David Murray, David Saunders, Jose Lourtie, Steffen Petersen, Mark Vertue, Joanne Riches, Julie Saunders, Mervyn Andiapen, Jonathan Breeze, Amy Hoare, Alison Pottle, Paula Rogers, Claire Prendergast, Michael Lewis, Gill Pout, Patricia Dickens, James Kirkby-Bott,

Pauline Bartlett, Guy Titley, Emma Willett, Nina Barratt, Tanuja Shah, Kathleen Holding, Lianne Hufton, Veeranna Shatkar, Ruwan Weerakkody, Caron Baldwin, Sarah Hare, Annette Woods, Ewen Griffiths, Arlo Whitehouse, Jugdeep Dhesi, Jane Okello, Philip Braude, Karen Wilson, Kirsty Gibson, Abdul Quddus, Davina Ross-Anderson, Katherine MacGloin, Hasan Mukhtar, Kathryn Simpson, Kayleigh Gilbert

Abstract

Introduction

Outcome of emergency admissions is usually limited to mortality with little attempt to capture the views of health status of survivors. This is because of the challenge of determining patient reported outcome measures (PROMs) for the period before their emergency admission. The aim was to assess the feasibility of collecting retrospective PROMs to capture the pre-admission health status of patients admitted as emergencies.

Methods

Prospective study of two cohorts: patients undergoing primary coronary angioplasty for acute ST-elevation myocardial infarction (STEMI) in five hospitals and emergency laparotomy (EL) for gastrointestinal conditions in 11 hospitals. Three rates were calculated: proportion of patients eligible for inclusion; proportion of eligible patients invited to participate; proportion of invitees who participated. Staff views were thematically analysed to understand factors that affected recruitment.

Results

About 85% of patients were eligible of whom most were invited to participate (84% EL; 79% STEMI). The proportions of invitees agreeing to participate differed between STEMI (92%) and EL (72%), probably reflecting greater post-intervention morbidity in the latter.

Variation between hospitals was observed in the proportion deemed eligible (EL 72-97%; STEMI 63-100%), proportion invited (EL 60-93%; STEMI 71-96%) and the proportion of invitees agreeing to participate (EL 55-92%; STEMI 67-100%). While this might reflect case-mix differences between hospitals, it suggests there is scope for less well-performing hospitals to improve their recruitment processes.

The extent to which this initial feasibility study was able to assess selection bias was limited to the age and sex of patients. There was no bias evident for EL patients but for STEMI, younger men were more likely to participate.

Conclusion

It appears to be feasible to collect retrospective PROMs from patients admitted unexpectedly as emergencies for the two conditions studied. The relevance of these findings to other causes of emergency admissions needs to be established. In addition, these findings justify the case for a large, multi-site study that could explore unresolved concerns about selection bias, particularly those arising from the clinical characteristics of patients. It would also enable estimates of the extent of variation in PROMs between hospitals to determine the usefulness of using PROMs in emergency admissions.

Introduction

In England, emergencies account for about 40% of all hospital admissions, with the number of admissions having increased by 47% over the last 15 years. Two-thirds of hospital beds are occupied by people admitted as emergencies and the cost is approximately £12.5 billion annually [1]. There is concern about variations in outcomes between providers [2][3][4]. While quite reasonably this has largely focused on mortality, there is also a need to consider outcome in terms of the health status of those who survive. To date, few attempts have been made to use patient reported outcome measures (PROMs) to determine patients' perception of any change in their health status. As the aim of healthcare is to restore a patient's health to his or her full potential, it is desirable to be able to compare patients' outcome with their health status before the sudden and unexpected event that led to an emergency admission. The use of PROMs would enable clinicians to review the impact of their care on individual patients and allow organisations, including regulators, to assess and compare the outcomes of different providers.

Using PROMs in emergency admissions presents the methodological challenge of how to capture a pre-event measure for such patients as pre-existing data are, inevitably, not available. A recent literature review [5] found strong agreement in elective patients between the PROM they reported before admission with their later recall of that pre-admission health status (via a retrospective PROM). This has been confirmed in England in a recent study of elective surgical patients [6]. These findings suggest that a retrospective PROM can provide a means of obtaining baseline health status in the absence of a prospectively collected contemporary report. Assuming this is also true for emergency patients (something that inevitably can never be established through direct testing), it is important to know whether it would be feasible to collect retrospective PROMs in such patients and the optimal methods for achieving this.

Feasibility might differ from the situation with elective admissions because, unlike elective admissions, emergency patients are acutely unwell and may be distressed. In addition, the immediate clinical priority is their surgical or medical assessment and intervention. Thus, it would not be possible to collect a PROM

until after initial treatment, during their recovery period some days later. Feasibility may also be influenced by the mode of administration and questionnaire design.

Only three studies have reported recruitment rates when using retrospective PROMs following emergency admissions. Two focused on trauma cases and one on acute lung injury. Gabbe and colleagues achieved 50% recruitment in trauma patients in two major Australian hospitals during their inpatient stay but boosted this to 77% by contacting them afterwards at home by mail and telephone [7]. Toien and colleagues who sought consent while trauma patients were in hospital in Norway and then surveyed them by mail afterwards achieved 50% recruitment [8]. Gifford and colleagues reported 70% recruitment among survivors of acute lung injury in four major hospitals in USA [9].

The aim of this exploratory study was to assess the feasibility of capturing retrospective PROMs in emergency admissions for a common medical (primary coronary angioplasty for acute ST-elevation myocardial infarction) and surgical (emergency laparotomy for gastrointestinal system) reason in a representative sample of NHS hospitals. The primary objectives were to explore the three stages of recruitment: the proportion of emergency admissions that were eligible for inclusion; the proportion of eligible patients who were invited to participate by staff; and the proportion of patients invited who participated. The secondary objectives were: to determine the representativeness of recruited patients as regards their age and sex; and to compare recruitment rates in different hospitals to determine the potential maximum rate obtainable and the associated organisational factors.

Methods

Choice of conditions

The two clinical conditions were selected as both are the subject of a national clinical audit which aims to collect detailed clinical data from all cases. The National Emergency Laparotomy Audit (NELA) includes all patients over the age of 17 years undergoing an emergency laparotomy for gastrointestinal conditions in NHS hospitals in England and Wales [10]. The Myocardial

Ischaemia National Audit Project (MINAP) collects data on all patients with acute ST-elevation myocardial infarction (STEMI) who undergo emergency primary coronary angioplasty [11][12]. Patients who met the national clinical audit criteria and were alive at discharge were considered for inclusion in this study. Patients were excluded if: they were not literate in English; judged not to have sufficient cognitive ability; or were not resident in the UK.

Design

A multi-site study was carried out to ensure there would be some variation in the specific organisation of patient recruitment and data collection. This would allow us to gain insights into the relative merits of recruiting in different settings and with different personnel involved [13]. For emergency laparotomy, 14 hospitals were selected on the basis of their high case ascertainment rates in NELA of which 13 agreed to participate. For STEMI, five primary angioplasty centres in London and the surrounding area were invited and all participated.

Sites were asked to recruit all eligible patients during a 15 week period. The study received ethics approval from South East Coast - Brighton & Sussex Research Ethics Committee (REC reference: 16/LO/2053) and it was incorporated in the NIHR Research Network Portfolio. Each site nominated someone to be the site lead (usually a consultant or senior research director) responsible for overseeing local data collection. Site leads then nominated study leads who undertook the data collection and liaised directly with one of us (EK) if any queries arose, completed a study log (see below), and stored and returned the data. At some sites, the site lead and study lead were the same person. Study leads at each site could delegate recruitment to appropriate members of the clinical team so the number of staff involved could vary.

Patient recruitment

Staff were provided with training in the form of video clips and written materials. These materials were developed by EK from prior experience of collecting retrospective PROMs for elective patients in two cohort studies [14]. Video materials were produced with the support of the University media team and research partners from the earlier study. EK also visited or held a telephone

conference with staff at each site prior to the start of data collection (Appendix: Study Flow diagram).

Patients were invited to participate once emergency medical and surgical treatments had been completed and as close to the discharge date as possible to ensure the immediate effects of the intervention (such as a general anaesthetic) were minimised. Clinical staff explained the study to patients and provided written information. Written consent was obtained from participating patients. Staff added a sticky label which included patients' NHS numbers and some socio-demographic data (date of birth, sex, address). A questionnaire was completed by recruited patients once during their inpatient stay. Those impeded by physical disability or sensory impairment could be assisted by staff or family members reading aloud the questions and/or recorded responses on the questionnaire. They were cautioned to avoid influencing the patients' views.

Study Log

Each study lead was required to complete a log covering every patient who met the national clinical audit criteria during the recruitment period. Staff recorded whether a patient met the eligibility criteria for inclusion in the study and if they were invited to participate. The date of consent of participants was also recorded. Patients' reasons for declining to participate were recorded if an explanation was offered without direct questioning.

Questionnaires

The questionnaires (paper hardcopy) included demographic information, self-reported co-morbidities, a disease-specific PROM and a generic PROM. The questionnaires contained instructions asking patients to recall how they were one month before their current admission. A systematic review identified suitable PROMs with adequate psychometric properties. Clinicians were then consulted in an unstructured meeting (a formal consensus development method was not used) to determine the final choice. This included consideration of the length and likely burden on patients of instruments.

For emergency laparotomy, the Gastrointestinal Quality of Life Index (GIQLI) developed by Eypasch and colleagues was selected [15]. It consists of 36 questions relating to the gastrointestinal system and the impact of symptoms

and treatment on individuals' physical, emotional and social status. It takes 5-10 minutes to complete and has good test-retest reliability (intra-class correlation coefficient = 0.92), and internal consistency (Cronbach's alpha >0.90). The GIQLI is the most commonly used validated PROM in studies investigating outcomes in emergency abdominal surgery [16].

For STEMI, the Seattle Angina Questionnaire (SAQ-7) is a 7 item health status measure for patients with coronary artery disease that has well-established validity, reliability, sensitivity to clinical change, and prognostic value [17,18]. Scores range from 0–100, where higher scores indicate fewer symptoms and higher health-related quality of life. SAQ-7 has good domain coverage (symptom burden, functional status, and quality of life), psychometric properties (validity, sensitivity), feasibility to implement (questionnaire length, language availability, and cost to implement), and clinical interpretability (knowledge of how to interpret scores in a clinically meaningful way) [19].

Both groups completed a generic PROM, the EQ-5D-3L. This has five items concerning the domains of mobility, usual activities, personal care, pain/discomfort and anxiety/depression. It takes up to five minutes to complete [18]. For each of these questions, the respondent chooses from three responses indicating the level of their function. A multi-attribute utility score where death and perfect health are represented by 0 and 1 are calculated [19]. Scores less than 0 are considered worse than death and 1 is the maximum score possible. The EQ-5D-3L was used rather than the EQ-5D-5L as the former is still the version used in the National PROMs Programme in England.

Analysis

Quantitative Analysis

For each condition, three rates were calculated: the proportion of all admissions with the condition that staff considered met the eligibility criteria for inclusion; the proportion of eligible patients invited to participate by staff; and the proportion invited who participated. In addition, the representativeness of those participating was assessed by comparison with all those included in the national clinical audit, though this was only possible for age and sex as clinical

data were not available. The performance of hospital sites was compared to establish the maximum possible rates that could be obtained.

Qualitative Analysis

At the completion of the study, information was sought from the site leads using a structured form administered by telephone interview or email. These observations, supplemented by a field diary kept by EK, were subjected to thematic analysis to identify the factors that facilitated and impaired patient participation to learn how data collection might be maximised.

Results

Quantitative Results

Emergency laparotomy

Of the 13 hospitals that agreed to participate, 11 collected data for the full 15-week duration of the study. Two hospitals stopped after one month due to local staff changes and their data are not included in the analyses. Those two hospitals were the only ones where the site lead was a non-clinical audit manager.

In all 11 participating hospitals the site lead was either a consultant surgeon or anaesthetist. They took responsibility for identifying patients from their NELA database and ward lists and provided oversight of the data collection. In nine sites the study lead was a nurse (usually a research portfolio nurse). They invited and consented patients, and undertook the data collection on weekdays. In the other two sites, doctors took on these tasks. Some sites had additional staff to support managing the study log, arranging paperwork and covering periods of leave.

During the recruitment period, 546 emergency laparotomy patients were admitted and survived to discharge, of which 466 (85%) were deemed eligible to participate (Figure 5-1). Of the 80 ineligible patients, 64 were considered to lack capacity to consent and complete a PROM and 16 were not literate in English.

Of the 466 eligible patients, 395 (85%) were invited to participate. The main reasons for not inviting patients was that the patient was discharged rapidly

(e.g. transfer to another hospital, self-discharge) or discharged at weekends when staff collecting data were not at work.

Of the 395 invited, 268 (72%) patients agreed to participate and completed a questionnaire. Of the 127 who declined to participate, the most common reason recorded by staff was that they were feeling too tired to complete the questionnaire.

There was some variation across the 11 sites. The proportion deemed eligible ranged from 72 to 97%, those invited from 60 to 93% and those agreeing to participate from 55 to 92% (Table 5-1). There was no consistent relationship between the three rates (Figure 5-3). Causes of low overall recruitment could be because eligible patients were not invited (hospital J) or patients declined such invitations (F). Those with the highest overall participation included the hospital with the highest proportion deemed eligible (L) and the one with the lowest eligible proportion (A).

Patients who participated were representative of all admissions as regards sex (male 47% v 48%) and age (median 66 v 67 years) [10].

Primary angioplasty for STEMI

All five sites participated for the full study duration of 15 weeks. The site leads in four hospitals were the hospital or cardiology department research manager or director. In the other hospital a nurse consultant was site lead. The study lead responsible for recruiting patients and collecting data during weekdays at all sites was a nurse (primarily research portfolio nurses, with the support of ward nurses). Some sites had additional administrative research staff to support managing the study log and arranging paperwork.

A total of 636 ST-elevation myocardial infarction patients meeting the MINAP criteria were admitted during the 15 week study period and survived to discharge (Figure 5-2). 547 patients (86%) met the study's inclusion criteria and were eligible for invitation. Ineligible patients included 47 who lacked sufficient cognitive capacity, 36 not literate in English and 7 had no UK residence.

Of the 547 eligible to participate, 432 (79%) were invited by staff to participate. The main reasons for not inviting patients was that the patient was discharged rapidly (e.g. transfer to another hospital, self-discharge) and those discharged at weekends or at night when staff collecting data were not at work. Of the 432 invited, 396 (92%) patients participated and completed a questionnaire. Of the 36 who declined to participate, most provided no reason.

There was some variation across the five sites. The proportion deemed eligible ranged from 63 to 100%, those invited from 69 to 96% and those agreeing to participate from 67 to 100% (Table 5-2). Unlike with EL, there was some consistency in the relationship between the rates for the three stages (Figure 5-4). In hospital Q with the lowest recruitment proportion (33%), the rates were poor for all three stages. In contrast, hospital R with the highest proportion recruited (96%) achieved this by success in all three stages.

Patients who participated were more likely to be male (79% v 72%) and slightly younger (median: males 60 v 63 years; females 67 v 71) than all those included in the national clinical audit [10].

Qualitative Results

Staff identified facilitators and obstacles at each stage of recruitment.

1. Identification of eligible patients

Staff found the identification of eligible EL patients easier if the site lead was also involved in the national clinical audit. Identification was facilitated by combining their NELA register, emergency theatre lists and consultants' knowledge of patients. This was easiest in sites with a real-time NELA register and electronic patient trackers. Similarly, for STEMI identification was aided by the existence of pathway activation records. Conversely, for EL the relocation of patients (such as from ITU to ward) could delay identification as a patient could be temporarily 'lost'. This was rarely a problem for STEMI As patients were admitted to a designated ward or coronary care unit and rarely moved to other locations.

2. Inviting patients to participate

Timely identification of patients and their location was crucial to enable study leads to invite patients. The main reason that patients were not invited was because of missing the target period of 1-2 days before discharge. This was a particular problem at weekends. As many STEMI admissions stayed less than 48 hours, patients admitted on a Friday would be discharged over the weekend and thus risked not being invited as study leads were not available. The site that managed to capture all patients (R) did not routinely discharge patients over the weekend. One proposed solution is to involve members of the 'on-call' clinical team at weekends.

An additional challenge with EL patients was predicting when this window of opportunity would occur or when discharge would occur as there was greater variation between patients. One way of coping with this with EL patients was for staff to invite them as soon as they felt there was an opportunity to speak to them, such as after stepping down from ITU to the ward.

3. Gaining agreement from patients to participate

Staff felt that patient participation was more likely if they were approached in an open and positive manner, explaining the purpose of the study clearly. Also, bringing in members of the clinical team directly involved in their care helped.

Patients' attitudes about the reasons for PROMs, their health status and the extent to which they had come to terms with their emergency admission were factors that affected their agreement to participate. Patients understood and welcomed the value of PROMs when their purpose was explained by engaged staff.

Most patients were glad to be asked for their views. The perceived time involved affected some decisions. STEMI patients welcomed the brevity of the questionnaire and while some EL patients initially perceived the questionnaire to be too long, once they had seen that the questions were straightforward to complete (closed rather than open), most agreed to participate.

The main reason patients declined was they did not feel well enough to complete a questionnaire. Acceptance was greater once patients had had time to come to terms with the significant medical events they had experienced. As

staff, for ethical reasons, were not able to revisit a declined invitation when a patient felt better, there was a delicate balance needed between waiting for the patient to be well enough and missing the opportunity, such that they were discharged home already. Given that the speed of recovery varied between patients, it was difficult to always make the best judgment. Staff tried to invite as close to discharge as possible even if that risked missing patients.

Discussion

Main findings

Patients can successfully be recruited to complete PROMs during their inpatient admission following significant emergency treatment (primary angioplasty and emergency laparotomy). Identification of relevant patients presented few difficulties, partly because the patients were also being included in a national clinical audit. It may prove to be more problematic if no such audit existed.

Of those patients admitted, 86% met the eligibility criteria to be invited to complete a PROM questionnaire about their pre-admission health status. Most of those deemed eligible were invited (85% emergency laparotomy; 79% STEMI). The main reasons for not inviting patients was that the patient was discharged rapidly (e.g. transfer to another hospital, self-discharge) or at weekends or out-of-hours when staff collecting data were not at work.

Agreement by patients to participate differed between the two conditions: 92% for STEMI patients but only 72% for emergency laparotomy. This probably reflected the greater post-intervention morbidity of the latter group. Despite the modest participation rate of laparotomy patients, they were representative of all such patients as regards age and sex, but may have differed in other respects (clinical severity, comorbidities etc.). The observation that STEMI participants were more likely to be male and younger is of minimal concern given the very high participation rate among this group.

There was variation in the eligibility, invitation and participation rates between the hospitals. While some of this might reflect case-mix differences between hospitals (e.g. English literacy), these differences suggest that there is scope for less well-performing hospitals to improve their recruitment processes. The

reason for low eligibility in some sites (72% for laparotomy in hospital A and 63% for STEMI in hospital Q) requires investigation to see if these rates are clinically justified. Similarly, low invitation rates (62% in hospital J and 69% in hospital N) suggest that overall recruitment could be enhanced given that other sites achieved invitation rates of over 90%.

The lower proportions of patients agreeing to participate in some hospitals (55% in hospital F for laparotomy and 67% in hospital Q for STEMI) may reflect case-mix differences but it might be because staff were less enthusiastic and effective in how they approached and invited patients.

Comparison with other studies

This is the first study in England to demonstrate the feasibility of collecting retrospective PROMs in emergency hospital admissions. The only previous studies to report on recruitment rates of emergency admissions involved either trauma patients [7, 8] or critical care survivors (not all of whom were emergency admissions to hospital) [9]. Despite all three studies being concentrated in only 1-4 sites, in our multi-site study we achieved similar proportions of admissions participating (49-62%) to those previously reported (50-77%).

Strengths and limitations

Its strengths are that it considered both a common medical and surgical reason for an emergency admission, included both a disease-specific and generic PROM that varied in length, and we observed the recruitment performance in a wide range of 16 hospitals which differed in terms of annual volume of cases, teaching status and geographical location.

Despite this, some caution is needed in interpreting the generalisability of the results. First, the hospitals that participated in the emergency laparotomy study were those that were achieving a high case ascertainment rate in the national clinical audit so may have characteristics and a culture that is more likely to support the collection of PROMs. As regards the STEMI sites, all five were located in and around London (for greater ease of access for the research team) so may differ from other parts of the country. This may explain why patients were slightly younger than that seen nationally.

Second, it is possible that the response of staff and patients to collecting PROMs for the two clinical conditions selected might not be replicated with other reasons for emergency admission. This will need to be investigated in subsequent implementation of retrospective PROM collection.

Third, about 14% of patients were excluded from this study as ineligible (those not literate in English and those with cognitive impairment). The proportion excluded varied between hospitals for EL (3-21%) and for STEMI (0-37%). Given that such differences could introduce some selection bias, future comparisons of hospitals' outcomes could be undermined if this was not taken into account. Investigations are needed to establish if such differences reflect the populations being served or the perceptions of staff as to the ability of patients to participate. In addition, recruitment of people not literate in English might be increased with the provision of questionnaires in other languages or translation services. For patients with cognitive impairment, the use of proxy-reported PROMs should be investigated.

Fourth, the extent of any selection bias was limited by the lack of data on patients' clinical characteristics. A further feasibility study with larger samples of patients, linked to their clinical characteristics, is needed so sub-group analyses could quantify the extent of any selection bias. This would also permit investigation of any desirability bias affecting which patients agree to participate.

Finally, given that the only means of obtaining patient reported health status in emergency admissions is by the use of retrospective PROMs, there will always be some uncertainty as to the impact of recall bias and response shift. However, unless it is believed that these biases differ systematically between hospitals, there is little risk to the meaningfulness of hospital comparisons.

Implications for practice

While the overall rates of eligibility, invitation and participation were good, they could be improved if those hospitals with lower rates adopted some of the processes that higher performing hospitals used. There are several potential ways of increasing recruitment in all three stages:

Support timely identification of patients

- integrate PROMs into the collection of data for the national clinical audit
- automatic PROM reminders as part of national clinical audit

Improving timing of invitation

- encourage the research nurses to participate in ward rounds to increase the support of ward nurses
- involve other clinical staff at weekends (if discharges fall on this day)

Improving staff ability to invite patients

- engage all relevant clinical staff to ensure the aim and purpose of collecting retrospective PROMs is understood
- embed PROMs collection with the completion of discharge forms
- reduce staff workload by simplifying the patient information sheets and consenting procedure.

Increasing patients' acceptance

- invite patients to participate as close to the discharge date as possible

Conclusions

This initial feasibility study has suggested that it is feasible to collect retrospective PROMs from patients admitted unexpectedly as emergencies for the two conditions studied across a variety of types of hospitals in the NHS. The relevance of these findings to other causes of emergency admissions needs to be established. In addition, these findings justify the case for a large, multi-site study that includes clinical information on participants, and could explore the unresolved concerns about selection bias, particularly those arising from the clinical characteristics of patients. It would also enable estimates of the extent of variation in PROMs between hospitals to determine the usefulness of using PROMs in emergency admissions.

Tables

Table 5-1 Emergency Laparotomy recruitment overall and by hospital (n=11)

	Hospital											
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**J**	**K**	**L**	**Overall**
N1 Number of admissions discharged alive	18	46	81	39	21	54	64	110	18	56	39	**546**
N2 Number of eligible patients	13	36	67	36	18	44	55	95	15	49	38	**466**
N3 Number invited to take part	12	31	62	27	15	40	49	80	9	42	28	**395**
N4 Number participated	11	20	42	20	11	21	30	50	8	33	22	**268**
N2/N1 Percentage of admissions deemed eligible	72	78	82	92	86	82	86	86	83	88	97	**85**
N3/N2 Percentage of eligible patients invited	92	86	93	75	83	91	89	84	60	86	74	**85**
N4/N2 Percentage of eligible patients participating.	85	56	63	56	61	47	55	53	48	67	59	**59**
N4/N3 Percentage of invited patients participating.	92	65	68	74	73	55	61	63	77	79	85	**72**
N4/N1 Percentage of admissions participating	61	43	52	51	52	39	47	45	44	59	56	**49**

Table 5-2 STEMI recruitment overall and by hospital (n=5)

	Hospital					
	M	**N**	**P**	**Q**	**R**	**Overall**
N1 Number of admissions discharged alive	180	156	123	49	128	**636**
N2 Number of eligible patients	152	129	107	31	128	**547**
N3 Number invited to take part	108	89	88	24	123	**432**
N4 Number participated	91	83	83	16	123	**396**
N2/N1 Percentage of admissions deemed eligible	84	83	87	63	100	**86**
N3/N2 Percentage of eligible patients invited	71	69	82	77	96	**79**
N4/N2 Percentage of eligible patients participating.	60	65	78	52	96	**72**
N4/N3 Percentage of invited patients participating.	84	93	94	67	100	**92**
N4/N1 Percentage of admissions participating	51	54	67	33	96	**62**

Figures

Figure 5-1 Recruitment Flow Diagram for Emergency Laparotomy patients

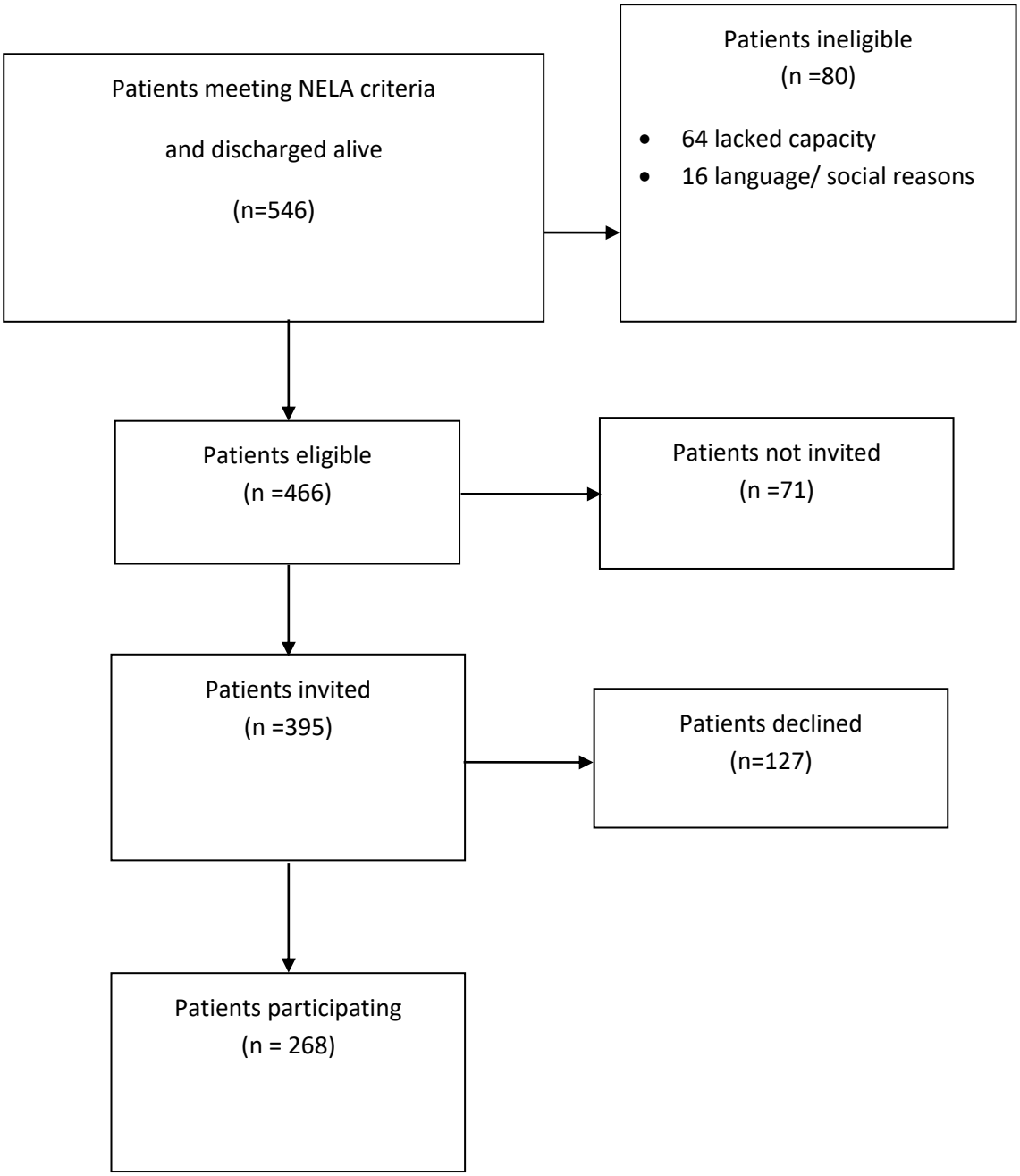


Figure 5-2 Recruitment Flow Diagram for STEMI patients

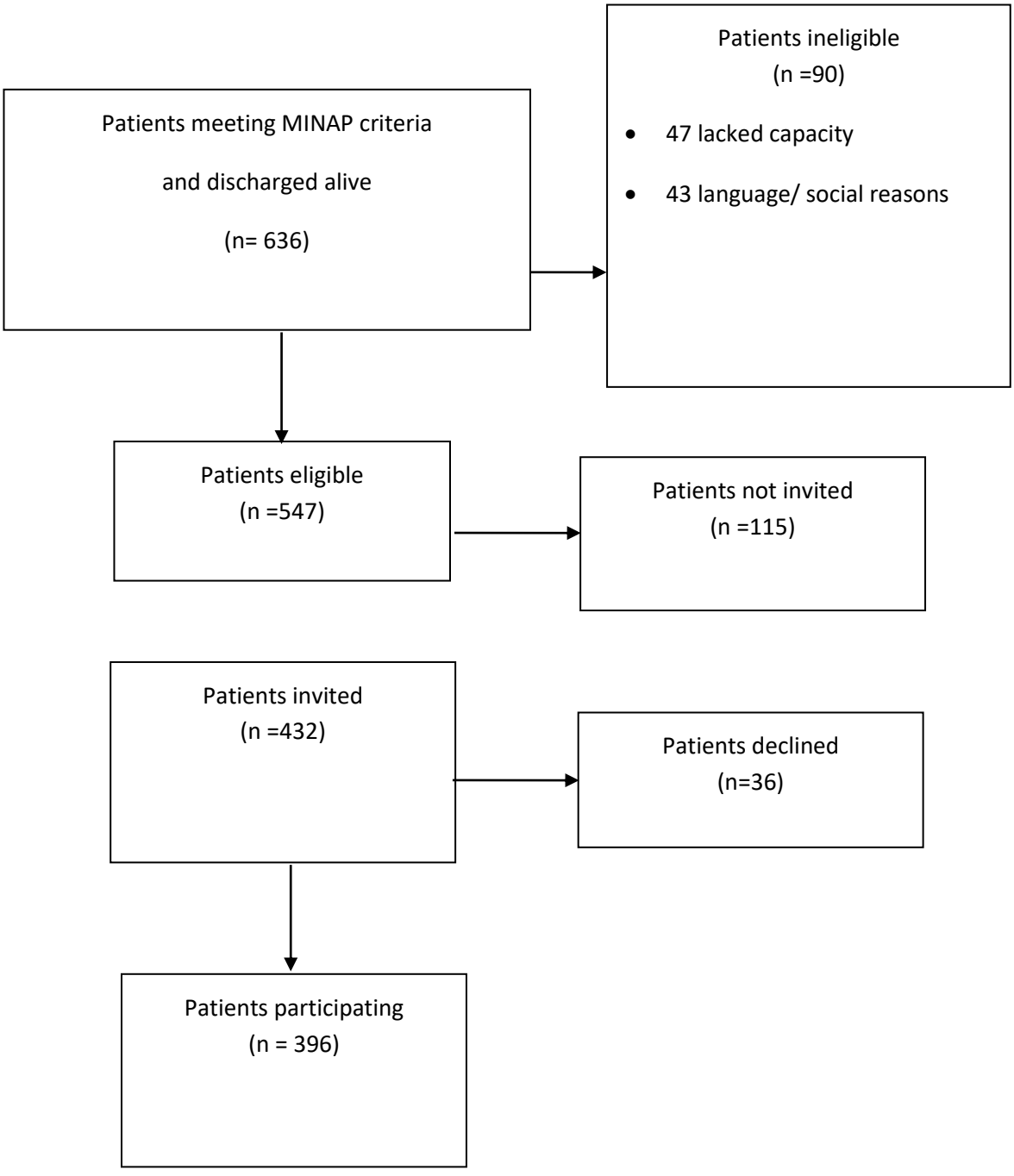


Figure 5-3 Relationship between the proportions of emergency laparotomy patients recruited at each of the three stages, by hospital

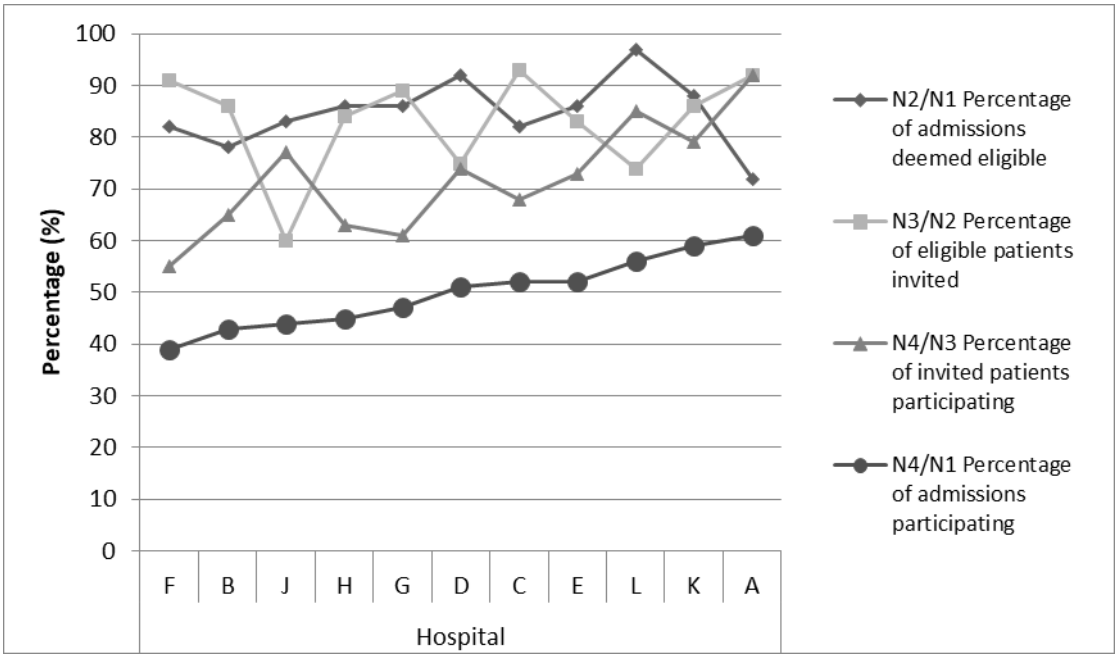
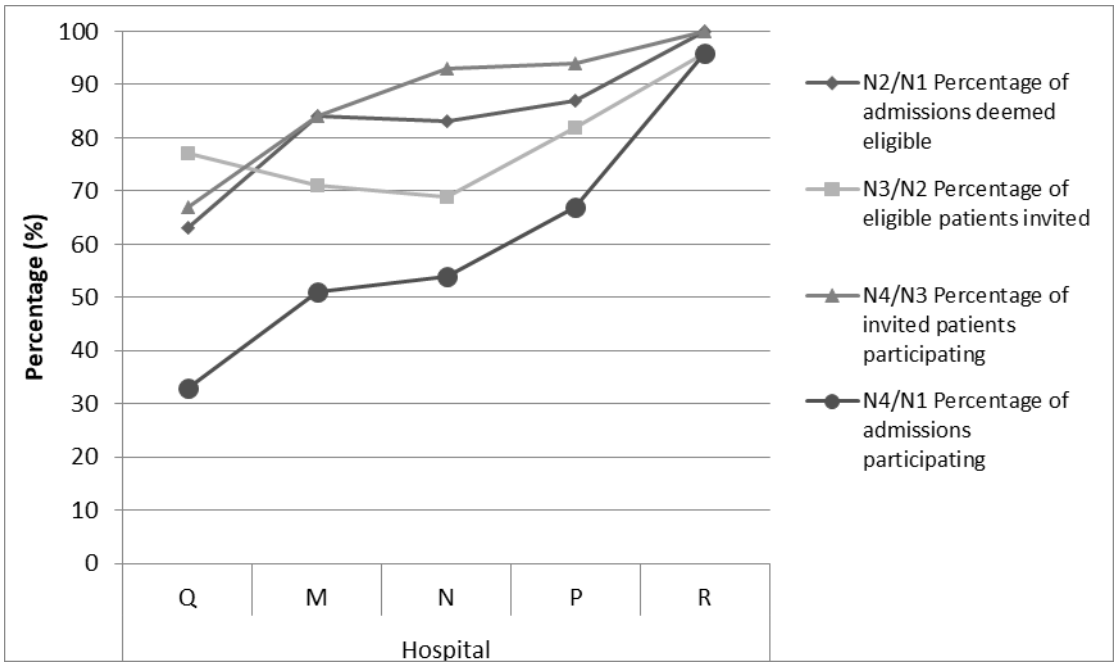


Figure 5-4 Relationship between the proportions of STEMI patients recruited at each of the three stages, by hospital



References

1 National Audit Office. Emergency admissions to hospital: managing the demand. *Hc* 2013;**739**:4.

2 NHS England. A&E Attendances and Emergency Admissions. May 2016 Monthly Report. https://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2016/06/Monthly-AE-Report-May-16.pdf. 2016.

3 Holt P, Sinha S, Ozdemir B, *et al.* Variations and inter-relationship in outcome from emergency admissions in England: a retrospective analysis of Hospital Episode Statistics from 2005-2010. *BMC Serv Res* 2014;**14: 270.**

4 The Royal College of Surgeons of England. Emergency surgery: standards for unscheduled care. 2011. https://www.rcseng.ac.uk/publications/docs/emergency-surgery-standards-forunscheduled-%0Acare%0A

5 Kwong E, Black N. Retrospectively patient-reported pre-event health status showed strong association and agreement with contemporaneous reports. *J Clin Epidemiol* 2017;**81**:22–32.

6 Kwong E, Neuburger J, Black N. Agreement between retrospectively and contemporaneously collected patient-reported outcome measures (PROMs) in hip and knee replacement patients. *Qual Life Res* Published Online First: February 2018.

7 Gabbe BJ, Cameron P a, Graves SE, *et al.* Preinjury status: are orthopaedic trauma patients different than the general population? *J Orthop Trauma* 2007;**21**:223–8.

8 Tøien K, Bredal IS, Skogstad L, *et al.* Health related quality of life in trauma patients. Data from a one-year follow up study compared with the general population. *Scand J Trauma Resusc Emerg Med* 2011;**19**:22.

9 Gifford J, Husain N, Dinglas V, *et al.* Baseline quality of life before

intensive care: A comparison of patient versus proxy responses. *Crit Care Med* 2010;**38**:855–60.

10 NELA Project Team. The second patient report of the National Emergency Laparotomy Audit. 2016;:1–160.http://www.nela.org.uk/reports

11 Chung S, Gale CP, Sundström J. Comparison of hospital variation in acute myocardial infarction care and outcome between Sweden and United Kingdom: population based cohort study using nationwide clinical registries. *BMJ* 2015;**351**:h3913.

12 National Institute for Clinical Outcomes Research, Herrett E, Smeeth L, *et al.* Myocardial Ischaemia National Audit Project. *Heart* 2010;**96**:1264–7.

13 Runkel PJ, McGrath JE. *Research on Human Behavior: A Systematic Guide to Method.* American Educational Research Association 1973.

14 Kwong E, Black N, J. Retrospectively collected patient reported outcomes show strong association and agreement with their contemporaneous reports. *J Clin Epidemiol* 2017;**81**:22–32.

15 Eypasch E, Williams JI, Wood-Dauphinee S, *et al.* Gastrointestinal Quality of Life Index: development, validation and application of a new instrument. *Br J Surg* 1995;**82**:216–22.

16 Stevens DJ, Blencowe NS, McElnay PJ, *et al.* A Systematic Review of Patient-reported Outcomes in Randomized Controlled Trials of Unplanned General Surgery. *World J Surg* 2016;**40**:267–76.

17 Spertus JA, Winder JA, Dewhurst TA, *et al.* Development and evaluation of the Seattle Angina questionnaire: a new functional status measure for coronary artery disease. *J Am Coll Cardiol* 1995;**25**:333–41.

18 Dougherty CM, Dewhurst T, Nichol WP, *et al.* Comparison of three quality of life instruments in stable angina pectoris: Seattle Angina Health Survey (SF-36), and Quality of Life Index. *Card Epidemiol*

1998;**51**:569–75.

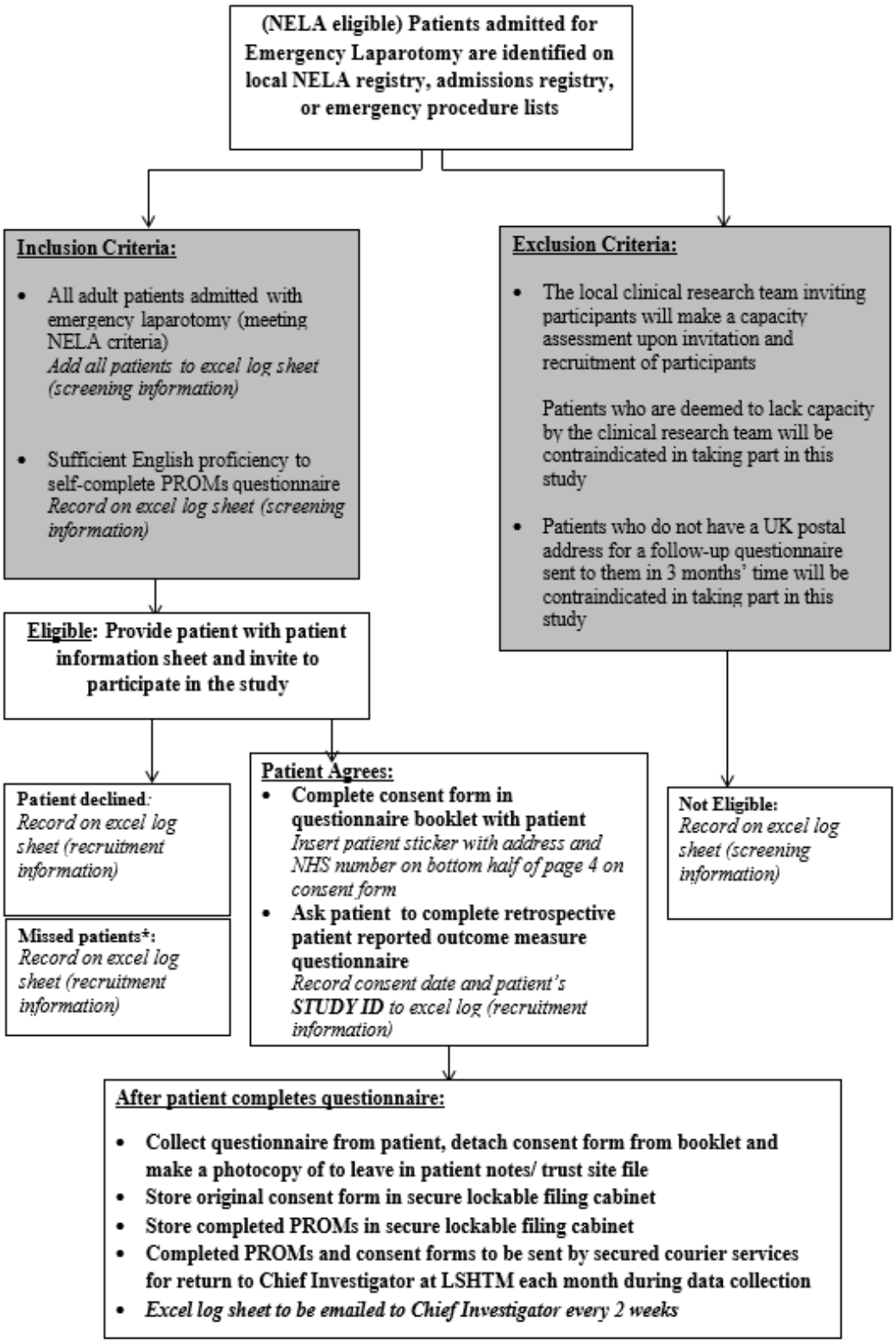
19 McNamara RL, Spatz ES, Kelley TA, *et al.* Standardized Outcome Measurement for Patients With Coronary Artery Disease: Consensus From the International Consortium for Health Outcomes Measurement (ICHOM). *J Am Hear Assoc* 2015;**4**.

20 EuroQol Group. EuroQol - a new facility for the measurement of health-related quality of life. *Health Policy (New York)* 1990;**16**:199–208.

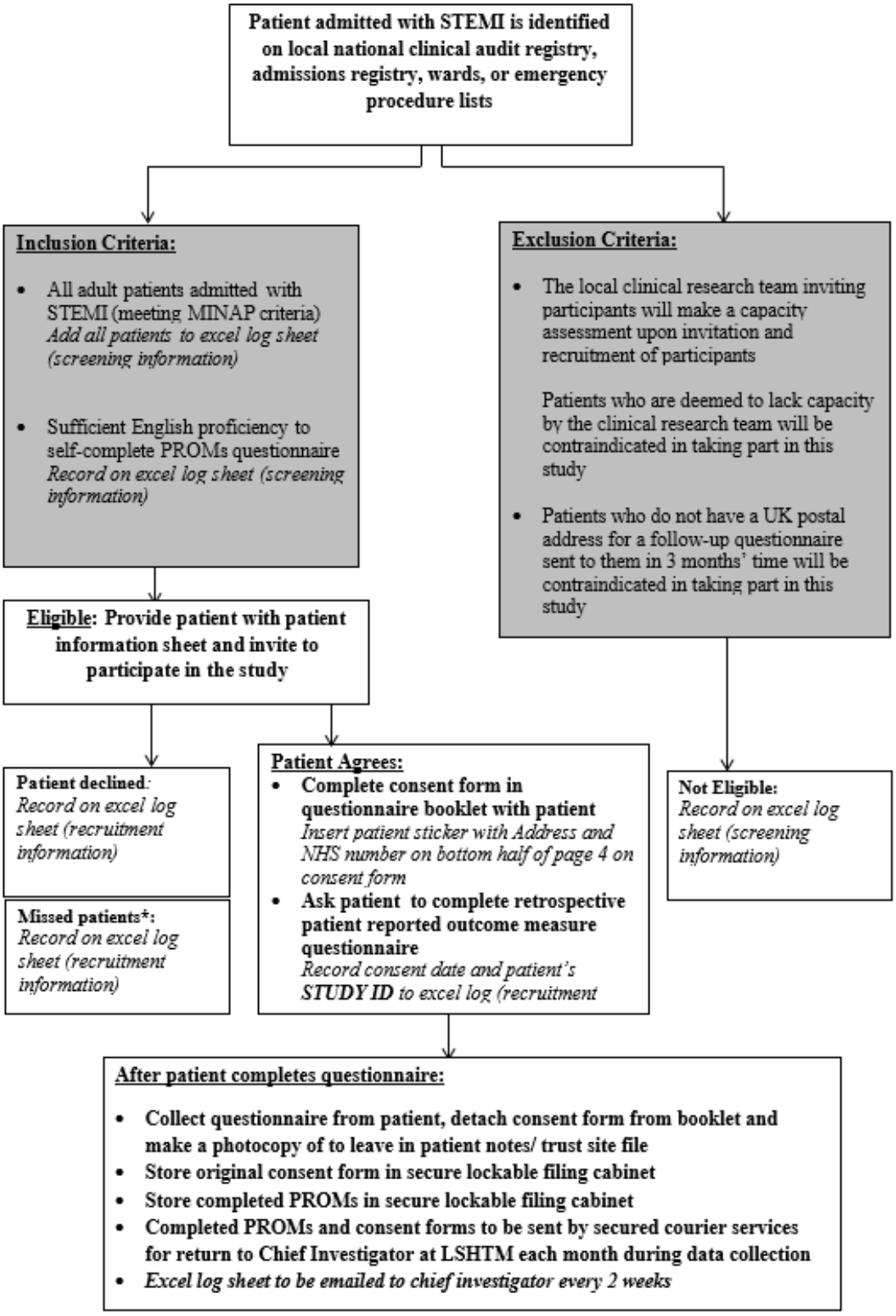
21 Dolan P, Roberts J. Modelling valuations for Eq-5d health states: an alternative model using differences in valuations. *Med Care* 2002;**40**:442–6.

Appendices

Emergency Laparotomy Study Flow Diagram



STEMI Study Flow Diagram



Chapter 6

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Esther Kwong
Principal Supervisor	Nick Black
Thesis Title	The use of patient reported outcome measures (PROMs) for evaluating emergency admissions

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	BMJ Open Gastroenterology
Please list the paper's authors in the intended authorship order:	Esther Kwong, Jenny Neuberger, Dave Murray, Nick Black
Stage of publication	**Submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	Please refer to details in the following page

Student Signature: [redacted] **Date:** 24 August 2018

Supervisor Signature: [redacted] **Date:** 24 August 2018

Chapter 6

Feasibility of collecting and assessing patient reported outcomes for emergency admissions: laparotomy for gastrointestinal conditions

After recruitment of patients described in Chapter 5, emergency laparotomy (EL) patients were followed-up with a contemporary PROM questionnaire to capture their outcomes.

This chapter describes the response rates achieved at three months when using PROMs in emergency laparotomy. It reports the patient characteristics of those who responded with those who did not respond, the assessment of the degree of any response biases and the interpretation of the health change between patients' baseline and outcome PROMs scores at three months. I conducted the follow-up study of mailed questionnaires by post at three months with the support of a part-time administrative assistant Mrs Christina Breach. I conducted data analysis with statistical advice from Dr Jenny Neuburger and prepared a manuscript with Professor Nick Black's guidance with regards to layout. I revised the manuscript following comments from co-authors. All co-authors approved the final version before journal submission. This manuscript has been submitted to BMJ Open Gastroenterology.

Feasibility of collecting and assessing patient reported outcomes for emergency admissions: laparotomy for gastrointestinal conditions

Esther Kwong

Research Fellow

Jenny Neuburger

Honorary Lecturer in Statistics

Dave Murray

Consultant Anaesthetist

Nick Black

Professor of Health Services Research

Authors Affiliation:

Corresponding Author:

Dr Esther Kwong, MBBS, MSc, BSc, Doctoral Candidate,

Dept. of Health Services Research & Policy,

London School of Hygiene and Tropical Medicine

15-17 Tavistock Place, London WC1H 9SH, UK.

Email: esther.kwong@lshtm.ac.uk Phone: +44 (0)207 958 8285

Dr Jenny Neuburger, Honorary Lecturer in Statistics, Dept. of Health Services Research & Policy, London School of Hygiene and Tropical Medicine

Dr Dave Murray, Consultant Anaesthetist, Anaesthetic Department, James Cook University Hospital, Middlesbrough, UK

Professor Sir Nick Black, Professor, Dept. of Health Services Research & Policy, London School of Hygiene and Tropical Medicine

Keywords

Patient Reported Outcome Measures, Health Status, Health-related quality of life, Retrospective, Response Rate Feasibility, Emergency admissions, Emergency Laparotomy

Contribution to the study

EK was the Principal Investigator of the study. NB is the Doctoral Supervisor. EK and NB wrote the paper with input from JN and DM.

Conflicts of Interest

We declare that we have no conflicts of interest.

Funding

EK is funded by Economic & Social Research Council doctoral fellowship. Grant Reference: ES/J500021/1. DM has been funded by NELA for his time as National Clinical Lead (2012-2017) and Chair of the Project Team (since 2017).

The research was supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care North Thames at Barts Health NHS Trust. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Ethical approval and Informed Consent

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and national research committee (NHS Health Research Authority) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. NHS ethical approval obtained from South East Coast - Brighton & Sussex Research Ethics Committee (REC reference: 16/LO/2053). Informed consent was obtained from all individual participants included in the study.

Acknowledgements

We thank: The patients and staff of 11 hospitals that participated in the study. NELA project team. Site and Study leads.

Collaborators: Mike Grocott, David Murray, David Saunders, Jose Lourtie, Michael Lewis, Gill Pout, Patricia Dickens, James Kirkby-Bott, Pauline Bartlett, Guy Titley, Emma Willett, Nina Barratt, Tanuja Shah, Kathleen Holding, Lianne Hufton, Veeranna Shatkar, Ruwan Weerakkody, Caron Baldwin, Sarah Hare, Annette Woods, Ewen Griffiths, Arlo Whitehouse, Jugdeep Dhesi, Jane Okello, Philip Braude, Karen Wilson, Kirsty Gibson, Abdul Quddus, Davina Ross-Anderson, Katherine MacGloin, Hasan Mukhtar, Kathryn Simpson, Kayleigh Gilbert.

Abstract

Introduction

Audit of emergency surgery is usually limited to immediate clinical outcomes relating to outcomes during the acute hospital episode with little attempt to capture patients' views of their longer-term outcomes. Our aim was to determine the response rate to patient reported outcome measures (PROMs) for patients who underwent an emergency laparotomy for gastrointestinal conditions, identify response bias and explore the feasibility of comparing outcomes with their prior health based on their recalled view collected during their admission.

Methods

Patients undergoing emergency laparotomy in 11 hospitals were recruited to complete a retrospective questionnaire containing the EQ-5D-3L and Gastrointestinal Quality of Life Index (GIQLI). Response rate for 3-month mailed follow-up questionnaire and potential response biases were assessed. Patients' outcomes were compared with their baseline using chi-squared and paired t-test to assess for differences.

Results

Of 255 patients contacted at three months, 190 (74.1%) responded. Responders were more likely to be older, female and more affluent. Patients' health improved significantly as regards the GIQLI (93.3 v 97.9; p=0.048) and the sub-scale on symptoms (51.9 v 59.6; p<0.001). No significant change in sub-scales on emotion or physical aspects or for overall health status (EQ-5D: 0.58 v 0.64; p=0.06). According to the social sub-scale patients had deteriorated (11.0 v 9.8; p<0.0006). Differences in change scores by patient characteristics were slight, suggesting minimal response bias.

Conclusion

This approach offers the opportunity for assessing the impact of treatment, from the patient's perspective and the potential to evaluate emergency laparotomy care using PROMs.

Summary Box

What is already known about this subject?

- Laparotomy is one of the commonest emergency surgical interventions with higher postoperative morbidity and mortality than elective procedures.
- In elective surgery these outcomes can be supplemented by Patient Reported Outcomes Measures (PROMs), but they have not been used routinely for emergency admissions.
- Whilst the feasibility of asking emergency laparotomy patients to recall their pre-admission health status has been demonstrated, their likelihood of responding to a mailed post-discharge questionnaire is unknown.

What are the new findings?

- PROMs can be successfully collected in patients three months after emergency laparotomy with a response rate of 74% using mailed follow-up.
- Most patients have not only regained their prior level of gastrointestinal health but their general health also improved.

How might it impact on clinical practice in the foreseeable future?

- PROMs offer the opportunity for routinely assessing the impact of treatment from the patient's perspective.
- Meaningful comparisons of surgeons and hospitals based on PROMs could be undertaken to supplement clinical measures such as mortality, morbidity and complications.

Introduction

In England, 40% of NHS hospital admissions are emergencies and the rate has been rising[1] [2]. Annually there are about 600,000 emergency admissions for general surgery, making up approximately half of all general surgical admissions [3]. Laparotomy is one of the commonest emergency surgical interventions with a higher postoperative morbidity and mortality than elective procedures [4].

If the aim of healthcare is to restore a patient to his or her full potential, we need to be able to compare patients' outcomes with their health status before the sudden and unexpected event that led to their emergency admission. Patient Reported Outcomes Measures (PROMs) are one of the ways to measure effectiveness and to determine the benefit of resources spent [5][6]. PROMs are self-reported questionnaires designed to be completed by patients to capture their health at specific points in time to detect a health change over a period. They are multi-dimensional measures which may cover symptoms, functional status or health-related quality of life (HRQL) [6].

It is known that short-term clinical outcomes, such as morbidity and mortality, following emergency surgical care vary significantly between hospitals [7][8]. In contrast, little is known about the longer-term health status of those who survive, the vast majority of patients. Capturing PROMs would provide an additional means of routinely assessing the effectiveness of emergency surgical care. Currently, we know little about whether PROMs for emergency surgery vary between hospitals and whether there is any unwarranted variation.

There is minimal existing research about the feasibility of collecting routine follow-up PROMs from patients who have completed a PROM during their in-patient episode. The relevance of the available evidence is unclear as studies either involved only a few centres or were restricted to protocol-driven intervention trials instead of routine use [9][10]. In addition, studies were mostly conducted in other countries so the results may not be applicable in England [11–18]. Response rates ranged between 51% and 71% for mailed questionnaires, and between 51% and 84% for interviewer administered questionnaires. The only attempt in England to collect PROMs in multiple sites

involved 28 major trauma centres and achieved about a 50% response rate using mailed or online follow-up at 6 months (personal communication: Antoinette Edwards).

To determine the feasibility of employing PROMs in emergency admissions, we undertook two exploratory studies, one on a medical condition and the other in surgery (emergency laparotomy). Patients' recollected state of health prior to their admission was collected shortly after their laparotomy but before discharge from hospital to provide a baseline assessment. We have already reported on the feasibility of recruiting a representative sample of patients [19].

This paper reports on the follow-up response rate for patients, identifies any response biases and explores the feasibility of comparing patients' outcome at three months with their retrospectively collected PROMs at baseline.

Methods

Site and patient recruitment

A multi-site study was carried out to ensure there would be variation in the administration of patient recruitment and data collection. This would allow us to gain insights into the relative merits of recruiting in different settings and with different personnel involved. Fourteen hospitals were selected, on the basis of their high case ascertainment rates in the National Emergency Laparotomy Audit (NELA), of which 13 agreed to participate and 11 successfully recruited patients for the 15 week duration of the study.

Patients who met the NELA inclusion criteria and were alive at discharge were eligible for inclusion in this study unless they were not literate in English, deemed not to have sufficient cognitive ability, or were not resident in the UK. For NELA, all patients over the age of 18 years, having a general surgical emergency laparotomy in all NHS hospitals in England and Wales are eligible for inclusion and are enrolled on a prospective basis into the audit. The inclusion criteria for the audit aims to include all emergency gastrointestinal procedures on the stomach, large and small bowel, for conditions such as perforation, bleeding, abdominal abscess or obstruction, via open or laparoscopic approaches. Emergency laparotomies following elective surgical

complications are also included. Patients requiring vascular surgery, gynaecological surgery, surgery on the renal tract, appendicectomy for appendicitis and laparotomy following trauma are excluded from the audit [20].

Patients were invited to participate after surgery, before discharge, and as close to the discharge date as possible to ensure the immediate effects of the intervention (such as a general anaesthetic and immediate post-operative complications including ileus, respiratory depression and side effects of opioids) were minimised to ensure that the patients were medically able to complete the questionnaire [20]. Clinical staff explained the study to patients, provided written information and obtained written consent. Questionnaires recalling their pre-admission baseline health status were completed by patients without assistance from staff or family except when they were impeded by physical disability or sensory impairment.

The study received ethics approval from South East Coast - Brighton & Sussex Research Ethics Committee (REC reference: 16/LO/2053) and it was incorporated in the NIHR Research Network Portfolio.

Three Month Follow-up

Patients were mailed a follow-up questionnaire (QF) from LSHTM 12 weeks (84 days) after their date of admission to hospital. Patient vital status was first checked against the Personal Demographics Service at NHS Digital prior to sending a follow-up questionnaire. After two weeks, non-responders were sent a reminder questionnaire.

Questionnaires

The questionnaires completed during the admission included demographic information, self-reported co-morbidities, a disease-specific PROM and a generic PROM. Patients were asked to recall how they were a month before their current admission. A systematic review identified suitable PROMs with adequate psychometric properties. Clinicians were then consulted in an unstructured meeting (a formal consensus development method was not used) to determine the final choice. This included consideration of the length and likely burden on patients of instruments.

The disease-specific PROM was the Gastro-Intestinal Quality of Life Index (GIQLI), developed by Eypasch and colleagues [21]. It consists of 36 questions relating to the gastrointestinal system and the impact of symptoms and treatment on individuals' physical, emotional and social status. It takes 5-10 minutes to complete and has good test-retest reliability (intra-class correlation coefficient 0.92), and internal consistency (Cronbach's alpha >0.90). The GIQLI is the most commonly used validated GI system specific PROM for studies investigating outcomes in emergency abdominal surgery The GIQLI score provides a global index score from 0 (poor health) to 144 (excellent health). The index score comprises four subscales: GIQLI symptoms (0-76), GIQLI physical score (0-28), social score (0-16) and emotion score (0-20) (Appendix 1). One item, on sex life, may not be applicable for some patients but the option of such a response is not available. Despite this, some patients wrote 'not applicable' on their questionnaire. They were coded as 'not at all'.

The generic PROM used was the EQ-5D-3L which has five items: mobility, usual activities, personal care, pain/discomfort and anxiety/depression. It takes up to five minutes to complete [22, 23] . For each of these questions, the respondent chooses from three responses indicating the level of their function. A multi-attribute utility score where death and perfect health are represented by 0 and 1 are calculated [23]. Scores less than 0 are considered worse than death and 1 is the maximum score possible. The EQ-5D-3L was used rather than the EQ-5D-5L as the former is still the version used in the National PROMs Programme in England.

Analysis

Participating patients' characteristics were summarised using means and SDs for continuous variables or percentages for binary variables. Response rates were calculated and reported for patients grouped by age, sex, living arrangements, socioeconomic status (SES), baseline GIQLI scores and baseline EQ-5D scores. SES was measured using the English Index of Multiple Deprivation (IMD) based on patients' residential postcodes (24) with patients assigned to quintiles of the national ranking of IMD scores.

We conducted chi-square and paired t-test for differences to compare characteristics of participants who responded to the 3-month follow-up questionnaire (QF) with those who did not. Patients' outcomes at 3 months were compared with their baseline using paired t-test to assess evidence of change in health status. Change scores, with the 95% confidence intervals, were also used to describe reasonable limits on the extent of any change, in order to assess whether the results were consistent with recovery to baseline (no change or an improvement in scores).

We explored the impact that non responses may have had on the mean health change in PROMs scores. To do this we calculated the change in score for patients by subset according to their characteristics (e.g. age, SES). We could then apply these estimates to the whole of the baseline cohort (regardless of whether they responded or not) to estimate what the mean change would have been if all had responded for the patient characteristics shown have a statistically significant non-response association. Inevitably such estimations assume that non-responders would have reported similar PROM changes as responders.

Results

Response rates

268 patients were recruited and completed baseline questionnaires (Appendix 2). Of these, 13 (4.9%) patients who were discharged from hospital then died during the post-discharge period before the follow-up contact. Of the 255 survivors, 190 patients (74.1%) responded to the follow-up PROM questionnaire: 146 responded to the first request and 44 after one reminder.

The mean time between completing the baseline (Q1) and the follow-up questionnaire (QF) was 85 (SD 19) days, and between admission and QF, 94 days.

Response bias

Responders and non-responders were similar as regards comorbidities, living arrangements and health status (EQ-5D and GIQLI) (Table 6-1). Responders

differed from non-responders in three ways: they were older (mean age 65.0 (SD 16; range 18-91) vs. 53.4 (SD 18; range 19-88) ($p<0.0001$) (Figure 6-1); more likely to be women; and more likely to come from more affluent SES.

Comparing change in PROM scores

The distribution of the EQ-5D at baseline was bimodal, with the majority of patients above 0.5 and a smaller peak between -0.5 and 0.5 (Figure 6-2). The distribution of the GIQLI score was broadly normal with a left skew.

Three months after surgery, patients' *GIQLI Emotion* score and *GIQLI Physical* score had returned to the baseline score (Table 6-2). The *GIQLI Symptoms* score had improved (51.9 v 59.6; $p<0.001$) whereas the *GIQLI Social* score had deteriorated (11.2 v 9.8; $p<0.001$).

The GIQLI score had improved (93.3 v 97.9, $p=0.048$) and EQ-5D score had improved considerably (0.58 v 0.64), although this difference was not statistically significant ($p=0.06$).

Influence of non-response on change in health status

Change in the GIQLI score and in the EQ-5D score was not associated with patients' SES (Table 6-3). However, change was greater in younger (under 70 years) and female patients though the differences did not reach statistical significance except for EQ-5D in women.

Assessment of non-response bias

Assessment of potential biases that might have been introduced by some patients not responding was based on the assumption that patients with similar baseline characteristics (sex and age) would have had similar follow-up EQ-5D or GIQLI scores. To illustrate the impact on non-response linked to sex and age, we estimated the mean change in GIQLI and EQ-5D scores had there been 100% follow-up response rate, compared to the observed mean changes. With this assumption, if responses were as per recruitment proportions by gender, the change in GIQLI would have been 4.55 (for all participants including non-responders) compared to 4.60 (observed in responders) and the

mean change in EQ-5D would have been 0.055 compared to the observed mean change of 0.060.

If responses were as per recruitment proportions by age, the change in GIQLI would have been 5.10 instead of 4.60, and the mean change in EQ-5D would have been 0.061 instead of 0.060.

Discussion

Main Findings

Retrospective and three-month follow-up PROMs can be successfully collected in representative samples of patients undergoing emergency laparotomy surgery across the country with a response rate of 74% using mailed follow-up. Although responders and non-responders were similar with regards to their living arrangements, number of co-morbidities and baseline health status, responders were more likely to be older, women and of a higher socioeconomic status. The impact of any response bias appears to be slight. Response bias due to sex could overestimate the improvement in health status by 1% (0.05/4.45) on the GIQLI score and by 9% (0.005/0.060) on the EQ-5D index. In contrast, age bias may underestimate the improvement by 10% (0.5/4.6) on the GIQLI score and by 2% (0.001/0.060) on the EQ-5D.

The mean GIQLI had improved by three months from 93.3 to 97.9. This suggests that not only do patients regain their prior level of GI health after major emergency surgery but there is an improvement compared with a month before their emergency admission. *GIQLI symptoms* also improve, by 8 when compared to baseline, though *GIQLI social* decreased by 1.3. Patients' overall health status measured by the EQ-5D showed a considerable increase (0.58 v 0.64) although this was not quite statistically significant.

What this study adds

This study has demonstrated the feasibility of collecting PROMs three months after emergency surgery among patients who, during their admission, had supplied retrospective accounts of the pre-event health status. It has shown that with high response rates, any responder bias is slight and will not undermine comparisons of providers.

The observation that the *GIQLI social score* worsens despite *the symptom* score improving was unexpected. It may be that the use of retrospective reporting of pre-operative symptoms exaggerates their severity though such a bias was not detected in studies of elective surgery [25,26]. It could be that the *GIQLI social score* items require a longer recovery trajectory than *GIQLI Symptom* items.

The improvement of generic health status, as seen by the increase in EQ-5D, may reflect that emergency laparotomies are primarily performed in lifesaving situations; the improved health outcomes would imply that not only are these procedures lifesaving and restorative but also goes further and improves the quality of life of patients. This is not unsurprising, as a proportion of emergency laparotomies will be performed for conditions that may be associated with chronic symptoms prior to acute presentation (such as acute colonic perforation in diverticular disease). As such, recall of symptoms in the month preceding surgery may also encompass the impact of chronic disease.

Strengths and Limitations

This is the first study of using retrospective PROMs to collect patients' baseline health status and a three-month follow-up for those admitted for emergency surgical operations in England. It was also conducted in multiple sites (11 hospital trusts) in different regions in England. This confirmed the feasibility of recruiting patients from diverse different geographical populations, as well as assessing PROMs use in different hospital organisational cultures and environments.

One limitation is that some patients did not respond to the GIQLI item on their sexual life as there was no option to report 'not applicable'. A second limitation was that some categories of patient (defined by their cognitive or literacy ability) were not eligible for inclusion in NELA so could not be included in this study.

Another potential limitation is that despite the left skew of the GIQLI and EQ-5D data, we opted to use the same statistical test (paired t-test) for comparisons between the three month and baseline data for three reasons.

First, it enabled preservation of consistency in our comparisons between all the measures. Second, the sample sizes satisfied guidelines for using parametric comparisons [27,28]. And third, the t-test does not require the assumption of equal dispersion (equal variance) in the data when comparing between groups. However, as the t-test does not fully take into account the skew and truncation of the EQ-5D data, the confidence interval is the more appropriate method of interpretation of any differences and the p value should be interpreted with caution.

One further limitation is that only one follow-up was conducted. Further follow-ups would provide insight into the recovery trajectory of emergency laparotomy patients.

Conclusion

This approach assesses from the patient's perspective, the impact of emergency laparotomy treatment. It also offers an insight into the opportunity for assessing other hospital admissions that are emergencies. The generalisability of these findings needs to be investigated with research on other causes of emergency admissions.

Further research is needed to explore longer-term outcomes enabling mapping of recovery trajectories. In addition, by capturing clinical data on patients (e.g. P- POSSUM scores), such as by linkage to national clinical audit data, it would be possible to determine any association with diagnosis and severity. This would be essential to be able to make meaningful comparisons of hospitals' outcomes and to ensure the PROMs data could support clinical decisions.

Routine collection of PROMs in emergency admissions could be feasible by their inclusion in national clinical audits. Such data would enhance quality improvement by including, alongside clinical outcomes, information on patients' views of their symptoms, functional status and quality of life. For patients undergoing emergency laparotomy, there is a paucity of information available on the longer-term functional outcomes. Evidence obtained from PROMs can help inform shared decision-making before undertaking potentially high-risk surgery.

Tables and Figures

Table 6-1 Characteristics of responders compared with non-responders

Patient characteristic	Overall (n=255)	Responders (n=189)	Non-responders (n=66)	p value*
Sex				
Male	118 (46.0)	80 (42.3)	38 (57.6)	0.03
Females	137 (54.0)	109 (57.7)	28 (42.4)	
SES				
1 (least deprived)	34 (14.8)	29 (17.1)	5 (8.3)	
2	47 (20.4)	37 (21.0)	10 (16.7)	
3	49 (23.3)	38 (22.3)	11 (18.3	0.03
4	49 (21.3)	37 (21.8)	12 (20.0)	
5 (most deprived)	51 (22.2)	29 (17.6)	22 (36.7)	
missing	25	19	6	
Comorbidities				
0	58 (24.2)	37 (21.2)	21 (33.3)	
1	78 (32.6)	64 (36.4)	14 (22.2)	
2	44 (18.4)	33 (18.7)	11 (17.5)	0.186
3	32 (13.4)	22 (12.5)	10 (15.9)	
4 or more	27 (11.3)	20 (11.4)	7 (11.1)	
missing	16	13	3	
Living arrangements				
With family	203 (79.6)	149 (82.3)	54 (84.3)	
Alone	40 (15.7)	30 (16.5)	10 (15.6)	0.685
Other	2 (0.8)	2 (1.1)	0	
missing	10	8	2	
Mean EQ-5D (SD)	0.57 (0.40)	0.58 (0.39)	0.54 (0.42)	0.494
missing	12	10	2	
Mean GIQLI (SD)	94.1 (31.3)	94.7 (31.4)	92.3 (31.04)	0.619
missing	25	18	7	

*from Chi-square

Table 6-2 Comparison of baseline and follow-up PROMs scores

PROM	Number with complete data	Baseline (Q1) Mean (SE, 95% CI)	Follow-up (QF) Mean (SE, 95% CI)	Change (95% CI, p value)
GIQLI	158	93.3 (2.55, 88.3-98.4)	97.9 (1.77, 94.4-101.4)	+4.6 (0.37 to 8.83, 0.048)
GIQLI symptom	168	52.0 (1.18, 49.6-54.2)	59.5 (0.76, 58.0-61.0)	+7.5 (5.68 to 9.32, <0.0001)
GIQLI emotion	177	12.0 (0.45, 11.12-12.9)	12.3 (0.35, 11.6-13.0)	+0.3 (-0.43 to 1.04, 0.37)
GIQLI physical	176	14.0 (0.61, 12.8-15.2)	13.3 (0.46, 12.4-14.2)	-0.7 (-1.68 to 0.28, 0.18)
GIQLI social	174	11.0 (0.34, 10.4-11.7)	9.8 (0.29, 9.27-10.4)	-1.2 (-1.82 to -0.58, 0.0006)
EQ-5D index	175	0.58 (0.03, 0.52-0.64)	0.64 (0.03, 0.59-0.69)	+0.06 (0.00 to 0.12, 0.06)

Table 6-3 Change in PROMs scores by age, sex and SES

Patient characteristic	Change in GIQLI (SD) (n=158)	p value*	Change in EQ-5D (SD) (n=160)	p value*
Age (years)				
>70	1.5 (25.0)		0.03 (0.36)	
50-70	6.9 (25.8)	0.46	0.07 (0.37)	0.70
<50	7.8 (39.3)		0.09 (0.50)	
Sex				
Male	2.46 (28.7)	0.43	-0.01 (0.40)	0.047
Females	6.14 (29.3)		0.11 (0.39)	
SES				
1 (least deprived)	2.39 (23.7)		0.13 (0.32)	
2	-0.75 (24.4)		0.47 (0.32)	
3	3.94 (28.1)	0.69	0.04 (0.39)	0.61
4	4.52 (26.4)		0.11 (0.42)	
5 (most deprived)	9.96 (27.5)		-0.01 (0.49)	

*from ANOVA

Figure 6-1 Age distribution of responders and non-responders

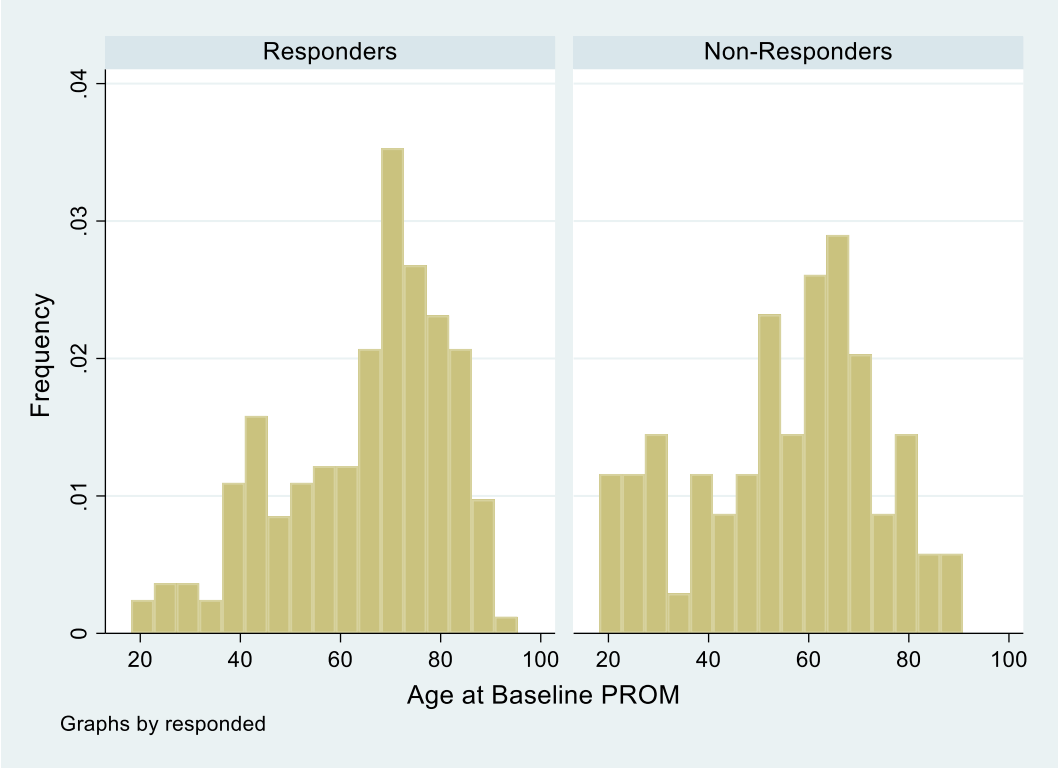
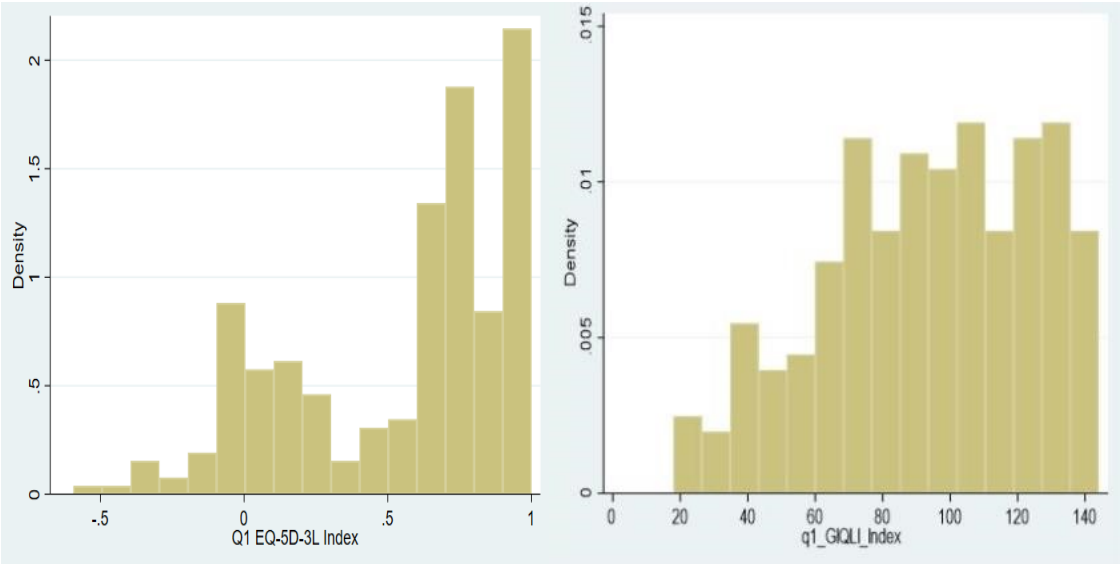


Figure 6-2 Baseline GIQLI score and baseline EQ-5D score distributions



References

1 NHS England. A&E Attendances and Emergency Admissions. May 2016 Monthly Report. https://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2016/06/Monthly-AE-Report-May-16.pdf. 2016.

2 Holt P, Sinha S, Ozdemir B, *et al.* Variations and inter-relationship in outcome from emergency admissions in England: a retrospective analysis of Hospital Episode Statistics from 2005-2010. *BMC Serv Res* 2014;**14: 270.**

3 The Royal College of Surgeons of England / Department of Health. The higher risk general surgical patient: Towards improved care for a forgotten group. 2011;:RCSENG-Professional Standards and Regulation.

4 Stoneham M, Murray D, Foss N. Emergency surgery: the big three-abdominal aortic aneurysm, laparotomy and hip fracture. *Anaesthesia* 2014;**69**:70–80.

5 Greenhalgh J, Dalkin S, Gooding K, *et al.* Functionality and feedback: a realist synthesis of the collation, interpretation and utilisation of patient-reported outcome measures data to improve patient care. *Heal Serv Deliv Res* 2017;**5**.

6 Black N. Patient reported outcome measures could help transform healthcare. *BMJ* 2013;**346**:19–21.

7 The Royal College of Surgeons of England. Emergency surgery: standards for unscheduled care. 2011. https://www.rcseng.ac.uk/publications/docs/emergency-surgery-standards-forunscheduled-%0Acare%0A

8 NELA Project Team. Third Patient Report of the National The third Patient Report of the National Emergency Laparotomy Audit (NELA). 2017.

9 Mason JD, Blencowe NS, McNair AG, *et al.* Investigating the collection and assessment of patient-reported outcome data amongst unplanned surgical hospital admissions: a feasibility study. *Pilot Feasibility Stud* 2015;**1**:16.

10 Parsons N, Griffin XL, Achten J, *et al.* Outcome assessment after hip fracture: is EQ-5D the answer? *Bone Joint Res* 2014;**3**:69–75.

11 Mock C, MacKenzie E, Jurkovich G, *et al.* Determinants of disability after lower extremity fracture. *J Trauma* 2000;**49**:1002–11.

12 Michaels a J, Madey SM, Krieg JC, *et al.* Traditional injury scoring

underestimates the relative consequences of orthopedic injury. *J Trauma* 2001;**50**:389–95; discussion 396.

13 Tidermark J, Zethraeus N, Svensson O, *et al.* Femoral neck fractures in the elderly: functional outcome and quality of life according to EuroQol. *Qual Life Res* 2002;**11**:473–81.

14 Watson WL, Ozanne-Smith J, Richardson J. Retrospective baseline measurement of self-reported health status and health-related quality of life versus population norms in the evaluation of post-injury losses. *Inj Prev* 2007;**13**:45–50.

15 Wilson R, Derrett S, Hansen P, *et al.* Retrospective evaluation versus population norms for the measurement of baseline health status. *Health Qual Life Outcomes* 2012;**10**:68.

16 Tøien K, Bredal IS, Skogstad L, *et al.* Health related quality of life in trauma patients. Data from a one-year follow up study compared with the general population. *Scand J Trauma Resusc Emerg Med* 2011;**19**:22.

17 Lyons R a, Kendrick D, Towner EM, *et al.* Measuring the population burden of injuries--implications for global and national estimates: a multi-centre prospective UK longitudinal study. *PLoS Med* 2011;**8**:e1001140.

18 Svedbom A, Borgström F, Hernlund E, *et al.* Quality of life after hip, vertebral, and distal forearm fragility fractures measured using the EQ-5D-3L, EQ-VAS, and time-trade-off: results from the ICUROS. *Qual Life Res* 2018;**27**:707–16.

19 Kwong E, Black N. Feasibility of collecting retrospective patient reported outcome measures (PROMs) in emergency hospital admissions. *Manuscr Submitt J Patient Rep Outcomes* 2018.

20 NELA. National Emergency Laparotomy Audit Inclusion Citeria. Published Online First: 2017.https://www.nela.org.uk/Criteria

21 Eypasch E, Williams JI, Wood-Dauphinee S, *et al.* Gastrointestinal Quality of Life Index: development, validation and application of a new instrument. *Br J Surg* 1995;**82**:216–22.

22 EuroQol Group. EuroQol - a new facility for the measurement of health-related quality of life. *Health Policy (New York)* 1990;**16**:199–208.

23 Dolan P, Roberts J. Modelling valuations for Eq-5d health states: an alternative model using differences in valuations. *Med Care* 2002;**40**:442–6.

24 Department for Communities and Local Government. English indices of deprivation. Crown Copyr. 2015. 2015.

25 Kwong E, Neuburger J, Black N. Agreement between retrospectively and contemporaneously collected patient-reported outcome measures (PROMs) in hip and knee replacement patients. *Qual Life Res* Published Online First: February 2018.

26 Kwong E, Black N. Retrospectively patient-reported pre-event health status showed strong association and agreement with contemporaneous reports. *J Clin Epidemiol* 2017;**81**:22–32.

27 Norman G, Streiner D. *PDQ Statistics*. 3rd ed. BC Becker Inc., Hamilton, London. 2003.

28 Minitab Blog Editor. Choosing Between a Nonparametric Test and a Parametric Test. [ONLINE]. Minitab Blog Ed. 2015.

Appendices

Appendix 1: GIQLI Questionnaire [21]

1. Over the past 2 weeks, how often have you had abdominal pains?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

2. Over the past 2 weeks, how often have you been bothered by a feeling of fullness in the upper abdomen?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

3. Over the past 2 weeks, how often have you felt bothered by bloating or the sensation of having too much gas in the abdomen?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

4. Over the past 2 weeks, how often have you felt bothered by passing wind?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

5. Over the past 2 weeks, how often have you felt bothered by burping or belching?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

6. Over the past 2 weeks, how often have you noticed unusual stomach or bowel noises?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

7. Over the past 2 weeks, how often have you been bothered by frequent bowel movements?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

8. Over the past 2 weeks, how often have you really enjoyed eating?

all the time,	most of the time,	now and then,	rarely,	never
(4)	(3)	(2)	(1)	(0)

9. How often have you had to refrain from eating the food you love due to your illness?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

10. Over the past 2 weeks, how did you manage to cope with everyday stress?

very badly,	badly,	moderately,	well,	very well
(0)	(1)	(2)	(3)	(4)

11. Over the past 2 weeks, how often have you felt sad about the fact that you are sick?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

12. Over the past 2 weeks, how often have you been nervous or anxious because of your illness?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

13. Over the past 2 weeks, how often have you been satisfied with your life in general?

all the time,	most of the time,	now and then,	rarely,	never
(4)	(3)	(2)	(1)	(0)

14. Over the past 2 weeks, how often have you felt frustrated about your illness?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

15. Over the past 2 weeks, how often have you felt tired or weary?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

16. Over the past 2 weeks, how often have you felt unwell?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

17. Over the past week (past 7 days), how many nights did you wake up at least once during the night?

every night,	5 to 6 nights,	3 to 4 nights,	1 to 2 nights,	never
(0)	(1)	(2)	(3)	(4)

18. To what extent has your illness led to disturbing changes in your appearance?

very much,	much,	somewhat,	a little,	not at all
(0)	(1)	(2)	(3)	(4)

19. To what extent has your general physical strength deteriorated due to your illness?

very much,	much,	moderately,	a little,	not at all
(0)	(1)	(2)	(3)	(4)

20. To what extent have you lost your stamina due to your illness?

very much,	much,	moderately,	a little,	not at all
(0)	(1)	(2)	(3)	(4)

21. To what extent have you lost your fitness due to your illness?

very much,	much,	moderately,	a little,	not at all
(0)	(1)	(2)	(3)	(4)

22. Over the past two weeks, have you been able to continue your normal daily activities (such as work, school, and household tasks)?

all the time,	most of the time,	now and then,	rarely,	never
(4)	(3)	(2)	(1)	(0)

23. Over the past 2 weeks, have you been able to continue your normal recreational activities (such as sports and hobbies)?

all the time,	most of the time,	now and then,	rarely,	never
(4)	(3)	(2)	(1)	(0)

24. Over the past 2 weeks, have you felt very restricted by the medical treatment?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

25. To what extent have your relationships with people close to you changed due to your illness?

very much,	much,	moderately,	a little,	not at all
(0)	(1)	(2)	(3)	(4)

26. To what extent has your sex life been impaired by your illness?

very much,	much,	moderately,	a little,	not at all
(0)	(1)	(2)	(3)	(4)

27. Over the past 2 weeks, have you been bothered by regurgitation of fluid or food?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

28. Over the past 2 weeks, how often have you felt bothered by your slow eating?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

29. Over the past 2 weeks, how often have you felt bothered by difficulty swallowing your food?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

30. Over the past two weeks, how often have you been bothered by urgent bowel movements?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

31. Over the past 2 weeks, how often have you been bothered by diarrhoea?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

32. Over the past 2 weeks, how often have you been bothered by constipation?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

33. Over the past 2 weeks, how often have you been bothered by nausea?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

34. Over the past 2 weeks, how often have you been alarmed by blood in your stool?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

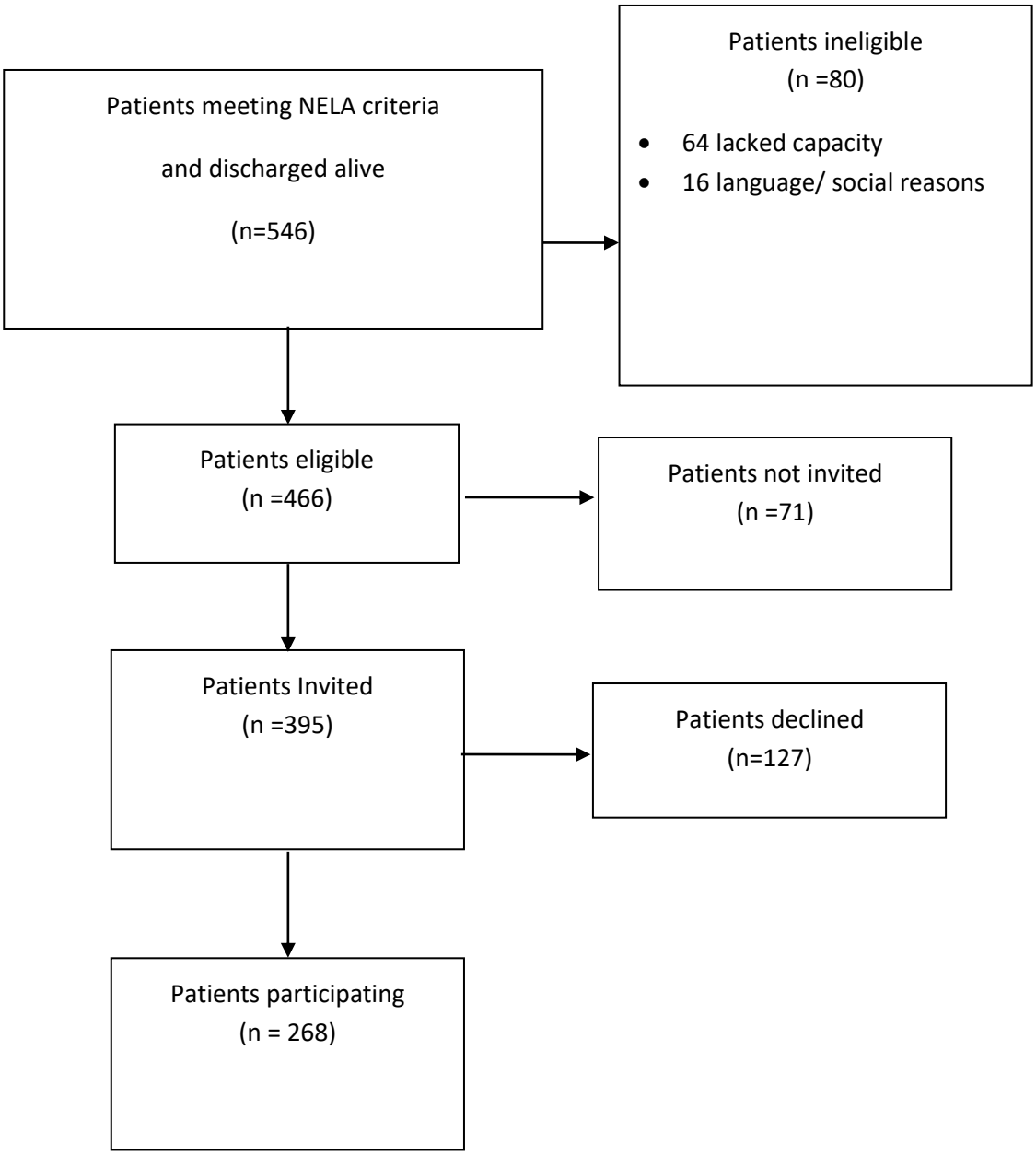
35. Over the past 2 weeks, how often have you been bothered by heartburn?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

36. Over the past 2 weeks, how often have you been bothered by involuntary bowel movements?

all the time,	most of the time,	now and then,	rarely,	never
(0)	(1)	(2)	(3)	(4)

Appendix 2: Study Flow Diagram



Chapter 7

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Esther Kwong
Principal Supervisor	Nick Black
Thesis Title	The use of patient reported outcome measures (PROMs) for evaluating emergency admissions

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	Open Heart
Please list the paper's authors in the intended authorship order:	Esther Kwong, Jenny Neuburger, Steffen Petersen, Nick Black
Stage of publication	**Submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	Please refer to details in the following page

Student Signature: [redacted] **Date:** 24 August 2018

Supervisor Signature: [redacted] **Date:** 24 August 2018

Chapter 7

Using patient reported outcome measures (PROMs) for primary percutaneous coronary intervention (PCI)

This chapter describes the response rates achieved at three months when collecting PROMs in STEMI patients, it reports on the patient characteristics of those who responded with those who did not respond, the assessment of the degree of any response biases and the interpretation of the health change between patients' baseline and outcome PROMs scores at three months. I conducted the follow-up study with PROMs questionnaires by mail at three months with the support of a part-time administrative assistant Mrs Christina Breach. I conducted data analysis with statistical advice from Dr Jenny Neuburger and prepared a manuscript first draft with Professor Nick Black's guidance on layout. I revised the manuscript following comments from co-authors. All co-authors approved the final version before journal submission. This manuscript has been submitted to Open Heart.

Using patient reported outcome measures (PROMs) for primary percutaneous coronary intervention

Esther Kwong

Research Fellow

Jenny Neuburger

Honorary Lecturer in Statistics

Steffen Petersen

Professor of Cardiovascular Medicine & Consultant Cardiologist

Nick Black

Professor of Health Services Research

Authors Affiliation:

Corresponding Author:

Dr Esther Kwong, MBBS, MSc, BSc, Doctoral Candidate,

Dept. of Health Services Research & Policy,

London School of Hygiene and Tropical Medicine

15-17 Tavistock Place, London WC1H 9SH, UK.

Email: esther.kwong@lshtm.ac.uk Phone: +44 (0)207 958 8285

Dr Jenny Neuburger, Honorary Lecturer in Statistics, Dept. of Health Services Research & Policy, London School of Hygiene and Tropical Medicine

Professor Steffen Petersen, Professor of Cardiovascular Medicine at the William Harvey Research Institute, Queen Mary University of London and a Consultant Cardiologist and Clinical Director for Research at Barts Heart Centre, Barts Health NHS Trust

Professor Sir Nick Black, Professor, Dept. of Health Services Research & Policy, London School of Hygiene and Tropical Medicine

Keywords

Patient Reported Outcome Measures, Health Status, Health-related quality of life, Retrospective, Response Rate, Feasibility, Emergency admissions, STEMI, AMI

Contribution to the study

EK was the Principal Investigator of the study. NB is the Doctoral Supervisor. EK and NB wrote the paper with input from JN and SP.

Conflicts of Interest

We declare that we have no conflicts of interest.

Funding

EK is funded by Economic & Social Research Council doctoral fellowship. Grant Reference: ES/J500021/1.

The research was supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care North Thames at Barts Health NHS Trust. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Ethical approval and Informed Consent

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and national research committee (NHS Health Research Authority) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. NHS ethical approval obtained from South East Coast - Brighton & Sussex Research Ethics Committee (REC reference: 16/LO/2053). Informed consent was obtained from all individual participants included in the study.

Acknowledgements

We thank: The patients and site leads and staff of 5 hospitals that participated in the study.

Collaborators: Julie Saunders, Mervyn Andiapen, Amy Richards, Mark Vertue, Joanne Riches, Jonathan Breeze, Amy Hoare, Alison Pottle, Paula Rogers, Claire Prendergast, Karen Wilson, Kirsty Gibson.

Abstract

Introduction

Routine measurement of the outcome of myocardial infarction is usually limited to immediate morbidity and mortality. Our aim was to determine the response to patient reported outcome measures (PROMs) three months later, identify response bias and explore the feasibility of comparing outcomes with their recalled view of their prior health state.

Methods

Patients admitted with STEMI to five PCI centres were invited to complete a retrospective questionnaire containing the EQ-5D-3L and Short Form Seattle Angina Questionnaire (SAQ-7). Response rate for a 3-month mailed follow-up questionnaire and potential response biases were assessed. Patients' outcomes were compared with their baseline using chi-square and paired t-test to assess for differences.

Results

Of 392 patients contacted, 260 (66.3%) responded. Responders were more likely to be older, female, more affluent, and have a higher EQ-5D at baseline. Three months after admission, patients' SAQ-7 and angina symptom subscale returned to their baseline score. The physical limitation subscale score was worse than at baseline (79.9 v 73.2, p=0.002), whereas the QoL subscale was better (66.6 v 73.9; p<0.001). The EQ-5D Index score was similar at 3 months and baseline (0.82 v 0.79). Evidence of bias arising from responders being in better general health at baseline needs further investigation and, if confirmed, needs to be taken into account in interpreting PROMs data.

Conclusion

It is feasible to use PROMs routinely to assess the impact of emergency admissions for STEMI patients. A larger demonstration project with more sites is needed to confirm these findings.

Summary Box

What is already known about this subject?

- While there have been improvements in the management of cardiovascular disease, significant variation still exists in survival following acute myocardial infarction (AMI) between hospitals within England.
- Morbidity and mortality outcomes can be supplemented by Patient Reported Outcomes Measures (PROMs), but have not been used widely in routine care.
- The feasibility of recruiting AMI patients to recall their pre-admission health status has been demonstrated, their likelihood of responding to a post-discharge mailed PROM questionnaire at three months is unknown.

What are the new findings?

- PROMs can be successfully collected in patients three months after STEMI with a response rate of 66.3% using mailed follow-up.
- Most patients regained their prior level of cardiac health as measured by the SAQ-7. The physical limitation subscale score was worse than at baseline, whereas, the QoL subscale was better.

How might it impact on clinical practice in the foreseeable future?

- PROMs offer the opportunity for routinely assessing the impact of treatment from the patient's perspective.
- Meaningful comparisons of hospitals based on PROMs could be undertaken to supplement clinical measures such as mortality, morbidity and complications.

Introduction

Despite the number of emergency admissions to hospital increasing and concern about variations in outcomes between providers [1,2] no attempt has been made to use patient reported outcome measures (PROMs) to determine patients' perception of their change in health status. In England, emergencies account for about 40% of all hospital admissions, with the number of admissions having increased by 47% over the last 15 years [3]. Two-thirds of hospital beds are occupied by emergencies and the cost to the NHS is approximately £12.5 billion annually [4].

Measuring the quality of healthcare is paramount for all health systems. PROMs is one of the ways to measure effectiveness and to determine the benefit of resources spent [5,6]. PROMs are self-reported questionnaires designed to be completed by patients to capture their health at specific points in time to detect a health change over a period. They are multi-dimensional measures which may cover symptoms, functional status or health-related quality of life (HRQL). Health status and quality of life are outcomes that are highly relevant and important to patients alongside traditional clinical outcomes and survival [6,7].

While there have been improvements in the management of cardiovascular disease, significant variation still exists in survival following acute myocardial infarction (AMI) between hospitals within England [8]. However, nothing is known about whether PROMs of survivors also vary between healthcare providers in England as routine assessment is limited to clinical outcomes (mortality and morbidity). Although there have been no attempts in England to routinely capture patients' recovery using PROMs they have been used in clinical trials. However, the extent to which such outcomes reflect those obtained in routine clinical care is unclear. There have been attempts to collect longer-term outcomes of AMI patients in the USA but whether those results are transferable to the English NHS is unclear [9–12].

If the aim of healthcare is to restore a patient to his or her full potential, we need to be able to compare patients' outcomes with their health status before the sudden and unexpected event that leads to the emergency admission. To

determine the feasibility of employing PROMs in emergency NHS admissions, an exploratory feasibility study was conducted in patients admitted with ST-segment elevation myocardial infarction (STEMI) for an emergency Primary Percutaneous Coronary Intervention (PCI) [13]. Success of recruiting patients soon after admission and of obtaining their recollected state of health prior to their admission to provide a baseline assessment has been reported [13–15].

In this paper we report on the follow-up response rate for patients who following an emergency admission, were confirmed to have suffered a STEMI, meeting the PPCI assessment checklist for inclusion and who underwent emergency (primary) percutaneous coronary intervention (PCI). Secondary objectives were to quantify any response bias as regards socio-demographic characteristics, comorbidity and health status and determine its potential impact on outcome assessment. Being an initial feasibility study, it was not powered sufficiently to make meaningful comparisons between participating centres.

Methods

Site and patient recruitment

A multi-site study was carried out to ensure there would be variation in the administration of patient recruitment and data collection. This would provide insights into the relative merits of recruiting in different settings and with different personnel involved [16]. For practical reasons, the study was confined to one region of England (North Thames). Five primary angioplasty centres were invited through the National Institute for Health Research (NIHR) Collaborations for Leadership in Applied Health Research and Care (CLAHRC) partnership network and all agreed to participate.

Patients admitted with STEMI for PCI to the five centres who were alive at discharge were eligible for inclusion unless they were: not literate in English; judged not to have sufficient cognitive ability; or were not resident in the UK.

Patients were invited to participate soon after their primary PCI and as close to the discharge date as possible to ensure the immediate effects of the intervention were minimised. Clinical staff explained the study to patients,

provided written information and obtained written consent. Questionnaires recalling their pre-admission baseline health status were completed by patients without assistance from staff or family except when they were impeded by physical disability or sensory impairment.

The study received ethics approval from South East Coast - Brighton & Sussex Research Ethics Committee (REC reference: 16/LO/2053) and it was incorporated in the NIHR Research Network Portfolio. Full details of the study methods and feasibility of recruitment have been described elsewhere [13].

Patients were sent a follow-up questionnaire by mail from LSHTM 12 weeks after their admission to hospital. Patient vital status was checked against the Personal Demographics Service at NHS Digital prior to sending a follow-up questionnaire. Non-responders after two weeks were sent a reminder questionnaire.

Questionnaires

The questionnaires completed during the admission included demographic information, self-reported co-morbidities, a disease-specific PROM and a generic PROM. Patients were asked to recall how they were one month before their admission.

The disease-specific PROM used was the short form Seattle Angina Questionnaire (SAQ-7 UK version). This is a 7 item health status measure for patients with coronary artery disease that has well-established validity, reliability, sensitivity to clinical change, and prognostic value [17–19]. Scores range from 0–100 (higher scores indicate fewer symptoms and higher health-related quality of life). SAQ-7 has good domain coverage (symptom burden, functional status, and quality of life), psychometric properties (validity, sensitivity), feasibility to implement (questionnaire length, language availability, and cost to implement), and clinical interpretability (knowledge of how to interpret scores in a clinically meaningful way) [20]. It assesses five dimensions: exertional capacity, angina stability, angina frequency, treatment satisfaction, and disease perception. Three sub-scales can be derived: physical limitation, angina symptoms, and quality of life (SAQ-QoL). The

summary scale and the three sub-scales extend from 0 (worst possible health state) to 100 (best possible health state). The SAQ-7 has been previously validated and applied in patients with acute coronary syndromes [17,18].

The generic PROM used was the EQ-5D-3L which has five items: mobility, usual activities, personal care, pain/discomfort and anxiety/depression. It takes up to five minutes to complete [21]. For each item, the patient chooses from three possible responses indicating the level of their function. A multi-attribute utility score where death and perfect health are represented by 0 and 1 is calculated [22]. Scores less than 0 are considered worse than death and 1 is the maximum score possible. The EQ-5D-3L was used rather than the EQ-5D-5L as the former is still the version used in the National PROMs Programme in England.

Analysis

Participating patients' characteristics were summarised using means and SDs for continuous variables, or percentages for categorical variables. Response rates were calculated and reported for patients grouped by age, sex, living arrangements, socioeconomic status (SES), baseline SAQ-7 scores and baseline EQ-5D scores. SES was measured using the English Index of Multiple Deprivation (IMD) based on patients' residential postcodes [23] with patients assigned to quintiles of the national ranking of IMD scores.

SAQ-7 scores and subscales were calculated according to scoring instructions from the questionnaire developers whereby partial responses were included where possible. Furthermore, individuals with non-responses to two or more items in a subscale did not contribute to the calculation of the component score as per scoring instructions provided by the developers of the SAQ-7 [17].

The likelihood of responding according to several patient characteristics (age, sex, SES, comorbidities, baseline SAQ-7 and EQ-5D) was calculated. This allowed the likely impact of non-response on the observed change in SAQ-7 and EQ-5D to be estimated for the patient characteristics shown to have a statistically significant non-response association, based on the assumption that non-responders would have reported similar PROM changes as responders.

Patients' outcomes at 3 months were compared with their baseline using chi-square and paired t-test to assess evidence of change in health status. Change scores, with the 95% confidence intervals, were also used to describe reasonable limits on the extent of any change, in order to assess whether the results were consistent with recovery to baseline (no change or an improvement in scores).

Results

Response rates

396 patients were recruited and completed questionnaires (Q1) recalling their health state one month earlier (Appendix: Study Flow Diagram). Of these, 4 (1%) died during the follow-up period. Of the 392 survivors, 260 patients (66.3%) responded to the follow-up PROM questionnaire (QF), 216 responded to the first request and 44 after the reminder.

The mean time between completing the baseline and the follow-up questionnaire was 89 (SD 17) days and between admission and follow-up questionnaire, 92 days.

Response bias

Responders and non-responders were similar as regards comorbidities, living arrangements and disease-specific PROM score (SAQ-7) (Table 7-1). Responders differed from non-responders in other ways: they were older (mean age 64.3 SD 12; range 35-94 vs. 57.1 SD 10; range 28-79, $p<0.0001$) (Figure 7-1); more likely to be women; more likely to come from more affluent SES; and have a higher generic PROM score (EQ-5D) at baseline.

Comparing change in PROM scores

The distribution of the EQ-5D at baseline has a left skew, with the majority of patients between 0.8-1.0 and a small minority having scores below 0.5 indicating poor health-related quality of life. The SAQ-7 score distribution at baseline also has a left skew but to a lesser extent than the EQ-5D index (Figure 7-2).

Three months after STEMI, patients' mean SAQ-7 score and mean angina symptom subscale was similar to their baseline score (Table 7-2). In contrast, the physical limitation subscale was worse than at baseline (79.9 v 73.2, p=0.002) while the SAQ-QoL subscale had improved (66.6 v 73.9; p<0.001). The EQ-5D Index score was slightly lower at 3 months than at baseline (0.82 v 0.79, p<0.02), although this is statistically significant, however, this appears to be due to a change in the shape of the distribution rather than a shift in distribution.

Influence of non-response on change in health status

Changes following STEMI and PCI in most PROMs scores (the SAQ-7, SAQ-7 subscales and EQ-5D) were not significantly associated with patient characteristics. The one exception was that patients in the poorest health (as determined by their baseline EQ-5D score), reported significantly larger (p<0.001) improvements in their EQ-5D scores at three months (Table 7-3).

Assessment of non-response bias

Assessment of potential biases that might have been introduced by some patients not responding was based on the assumption that patients with similar baseline EQ-5D index scores would have had similar follow-up EQ-5D or SAQ scores. To illustrate the impact on non-response linked to baseline EQ-5D (mean 0.82 in responders v 0.73 in non-responders, Table 7-1), we estimated the mean change in SAQ and EQ-5D scores had there been 100% follow-up response rate, compared to the observed mean changes. The mean change in SAQ-7 would have been 1.2 (estimated for all participants including non-responders) compared to 0.8 (observed in responders). The estimated mean change in EQ-5D would have been -0.02 compared to the observed mean change of -0.03.

Discussion

Main findings

Three-month follow-up PROMs can be successfully collected from two-thirds of patients admitted as emergencies with STEMI for primary PCI through mailed questionnaires. Although responders and non-responders were similar with regards to their living arrangements, number of co-morbidities and baseline

SAQ-7, responders were more likely to be older, women, of a higher socioeconomic status and be in better general health according to the EQ-5D score. Apart from the latter, none of these characteristics were associated with the change in health reported at follow-up so have not introduced bias to the findings. However, the higher EQ-5D of responders at baseline could introduce some bias leading to an underestimation of the improvement in the cardiac health of patients three months after the event: change in SAQ-7 would be 1.2 instead of 0.8. Similarly, the observed deterioration in generic health (EQ-5D -0.03 at follow-up) would be less (-0.02 at follow-up).

Three months after PCI, patients' mean SAQ-7 score and angina symptom score returned to their baseline score suggesting patients regain their prior level of cardiac health. Although patients reported greater physical limitation than beforehand, they felt their quality of life had improved. Given that a clinically meaningful difference in SAQ scores is estimated to be 5-8 points, these differences are clinically important [24].

Although the EQ-5D index score was lower at 3 months when compared to baseline (0.82 v 0.79), the clinical significance of this decrement should be further explored as although this reached statistical significance, the 95% confidence intervals overlap.

What this study adds

This study has demonstrated the feasibility of collecting PROMs three months after STEMI among patients who, during their admission, had supplied retrospective accounts of the pre-event health status. It has shown that the response is subject to responder bias which, if confirmed in a larger study, would need to be taken into account when comparing the outcomes of different providers to ensure meaningful findings.

The observation that whilst patients' physical limitation worsens, their quality of life (QoL) improves is surprising. There are four possible explanations. First, it may be that patients recall their prior QoL as worse than it was, although no such bias was detected in studies of elective surgery when retrospective and contemporary reports were compared [14,15]. Second, it may be that patients'

baseline disease-specific quality of life was already lowered due to the presence of sub-acute symptoms prior to their AMI, but were not at the clinical threshold that warranted medical attention. Grodzinsky et al reported similar baseline SAQ-QoL scores (63.8) in patients with AMI as that reported in this study [25]. Third, it may be that the physical limitation is due not to anginal symptoms measured by the SAQ-QoL score but by other complications following AMI or hospitalisation. Shortness of breath secondary to heart failure is a known complication following AMI which could limit patients' physical function [26,27].

And fourth, it may be that patients exercised caution in their physical exertion and hence imposed greater physical limitations on their function than necessary. Meanwhile, their QoL may improve from the psychological boost of having survived their AMI and had the reassurance of having had intervention for their coronary arteries. This may be due to a degree of response shift occurring following patients' experience of an AMI. Patients may have a different appreciation of their cardiac related quality of life. All PROMs, which are subjective reports, can be influenced by response shift [28]. The literature on clinical recovery trajectories after STEMI at three months is sparse with no studies reporting the SAQ-7. However, it is reported that patients continue to recover and improve their SAQ scores for up to 12 months [9,25].

The observation for the lower generic health status score (as measure by the EQ-5D) but an improvement in the SAQ-QoL score for patients at three months suggests that the former captures dimensions that may be important for STEMI patients that are non-disease specific. For example, there is evidence that depression following AMI is common [29], an aspect that is captured by the EQ-5D in its anxiety/ depression item but not by the SAQ-7.

Strengths and Limitations

This is the first study of using retrospective PROMs in routine clinical practice to collect patients' baseline and three-month health status for those admitted with STEMI in England. Conducting it in five NHS trusts demonstrated the feasibility of PROMs use in different hospital organisational cultures and environments.

One potential limitation is that despite the left skew of the EQ-5D data, we opted to use the same statistical test (paired t-test) for comparisons between the three month and baseline for three reasons. First it enabled preservation of consistency comparisons between all the measures. Second, the sample sizes satisfied guidelines for using parametric comparisons [30,31]. And third, the t-test does not require the assumption of equal dispersion (equal variance) in the data when comparing between groups. However, as the t-test does not fully take into account the skew and truncation of the EQ-5D data, the confidence interval is the more appropriate method of interpretation of any differences and the p value should be interpreted with caution.

A further limitation is that although the SAQ-7 has been validated for use in acute coronary syndromes [11], it is not specific for myocardial infarction and may not adequately capture some post-acute MI complications such as heart failure [17]. Further development of routine use of PROMs in emergency PCI patients should consider alternatives that were not available at the outset of this study (e.g. CROQ). Also, the addition of a PROM to capture post MI complications such as the Rose Dyspnoea Questionnaire may improve prognostic abilities in the evaluation of health change for patients after STEMI.

Implications for further research/ policy

This study shows that it is feasible to collect retrospective and follow-up PROMs from patients admitted as emergencies with STEMI in NHS hospitals. This approach offers an insight into the opportunity for assessing, from the patient's perspective, the impact of treatment for the 40% of hospital admissions that are emergencies, and patients' subsequent recovery after their emergency admission. The generalisability of these findings to other causes of emergency admissions needs to be established.

Further research is warranted to explore longer-term outcomes and compare these with patient risk profiles, clinical characteristics and recovery trajectories.

Routine collection of PROMs in emergency admissions is feasible using the retrospective PROMs collected during the index admission and a subsequent follow-up. Data could be linked to clinical measures known to be associated

with outcome (such as Kilip classification, concentration of Troponin I, infarct site and left ventricular ejection fraction) and quality dashboards to support on-going quality improvement through benchmarking, by promoting clinical effectiveness and patient-centred care. Larger studies are needed to collect PROMs in patients admitted with AMI and other emergency acute coronary syndrome patients to enable sub-group analysis of patient and clinical characteristics, to investigate further any response bias and to develop risk adjustment models to enable comparisons of providers.

Tables and Figures

Table 7-1 Characteristics of responders (n=260) compared with non-responders (n=132)

Patient characteristic	Overall Number (%)	Responders Number (%)	Non-responders Number (%)	p value
Sex				
Male	308 (78.6)	196 (75.4)	112 (84.9)	0.031
Females	84 (21.4)	64 (24.6)	20 (15.2)	
SES				
1 (least deprived)	68 (18.4)	48 (19.4)	20 (16.4)	0.013
2	60 (16.3)	50 (20.2)	10 (8.20)	
3	91 (24.7)	59 (23.9)	32 (26.2)	
4	94 (25.5)	53 (21.4)	41 (33.6)	
5 (most deprived)	56 (15.2)	37 (15.0)	19 (15.6)	
missing	23	13	10	
Comorbidities				
0	57 (14.6)	35 (13.6)	22 (16.8)	0.904
1	111 (28.5)	76 (29.5)	35 (26.7)	
2	94 (24.2)	64 (24.8)	30 (22.9)	
3	60 (15.4)	39 (15.1)	21 (16.0)	
4 or more	67 (17.2)	44 (17.1)	23 (17.6)	
missing	3	2	1	
Living arrangements				
With family	296 (75.9)	195 (75.2)	101 (77.1)	0.755
Alone	87 (22.3)	60 (23.2)	27 (20.6)	
Other	7 (1.79)	4 (1.54)	3 (2.29)	
missing	2	1	1	
Mean EQ-5D (SD)	0.79 (0.28)	0.82 (0.25)	0.73 (0.34)	0.002
Mean SAQ-7 (SD)	76.8 (21.1)	77.8 (22.3)	74.9 (20.4)	0.207

Table 7-2 Comparison of baseline and follow-up PROMs scores

PROM	Number with complete data	Baseline Mean (SE, 95% CI)	Follow-up Mean (SE, 95% CI)	Change (95% CI, p value)
SAQ_7 Summary	259	77.8 (1.27, 75.3-80.3)	78.6 (1.22, 76.2-81.03)	+0.8 (-1.6 to 3.2, 0.56)
SAQ_7 Physical Limitation	227	79.9 (1.9, 76.2-83.7)	73.2 (1.81, 69.6-76.8)	-6.7 (-10.3 to -3.1, 0.0018)
SAQ_7 Angina Symptom	258	86.9 (1.2, 84.6-89.2)	88.6 (14.1, 86.5-90.7)	+1.7 (-13.3 to 16.7, 0.24)
SAQ_QoL	254	66.6 (1.8, 63.0-70.2)	73.9 (1.7, 70.6-77.2)	+7.3 (3.9 to 10.7, <0.001)
EQ-5D index	256	0.82 (0.02, 0.79-0.85)	0.79 (0.02, 0.76-0.82)	-0.03 (-0.07 to 0.01, 0.02)

Table 7-3 Exploring the extent of differences in PROMs scores and health change with responder characteristics

Patient Characteristic	Mean SAQ-7 Summary Score Baseline (Q1) Score, (SD) (n=390)	Mean SAQ-7 Summary Score Follow-up (QF) Score, (SD) (n=260)	SAQ-7 Summary Score Difference in Health Change, (SD) (n=259)	p value* *from ANOVA
Age				
>70	80.3 (20.5)	78.4 (20.2)	-1.88 (25.0)	0.33
50-70	75.5 (21.0)	78.7 (19.0)	1.72 (22.3)	
<50	75.4 (21.8)	78.9 (20.7)	4.65 (25.5)	
Sex				
Male	77.3 (20.7)	80.4 (19.1)	1.92 (23.2)	0.20
Females	75.2 (22.4)	72.5 (20.7)	-2.46 (24.4)	
SES				
1 (least deprived)	79.6 (20.0)	83.4 (19.0)	5.89 (20.4)	0.23
2	80.0 (18.4)	79.5 (18.6)	-0.96 (23.8)	
3	78.5 (19.0)	77.2 (19.6)	-1.73 (23.0)	
4	72.6 (24.3)	80.4 (21.5)	5.10 (24.5)	
5 (most deprived)	75.1 (20.9)	72.6 (18.8)	-2.85 (27.9)	
EQ-5D baseline Categories				
1 (≤0.65)	59.1 (24.4)	64.5 (23.0)	6.3 (33.8)	0.16
2 (0.66-0.85)	74.7 (19.2)	76.4 (18.9)	2.4 (24.1)	
3 (0.86-1)	85.5 (15.9)	84.1 (16.7)	-1.6 (19.6)	

(Table continues overleaf)

Patient Characteristic	EQ-5D Baseline Score (n=385)	EQ-5D Follow-up Score (n=258)	EQ-5D Difference in Health Change (n=256)	p value* *from ANOVA
Age				
>70	0.80 (0.23)	0.78 (0.26)	-0.01 (0.24)	0.41
50-70	0.80 (0.28)	0.80 (0.22)	-0.05 (0.23)	
<50	0.76 (0.36)	0.74 (0.33)	-0.04 (0.23)	
Sex				
Male	0.80 (0.28)	0.81 (0.24)	-0.04 (0.23)	0.81
Females	0.75 (0.30)	0.71 (0.28)	-0.03 (0.23)	
SES				
1 (least deprived)	0.85 (0.21)	0.81 (0.23)	-0.04 (0.17)	0.38
2	0.83 (0.23)	0.84 (0.19)	-0.01 (0.26)	
3	0.79 (0.23)	0.80 (0.24)	-0.01 (0.23)	
4	0.73 (0.37)	0.78 (0.25)	-0.02 (0.24)	
5 (most deprived)	0.77 (0.34)	0.70 (0.31)	-0.10 (0.22)	
EQ-5D baseline Categories				
1 (≤0.65)	0.27 (0.28)	0.49 (0.38)	0.19 (0.33)	<0.001
2 (0.66-0.85)	0.78 (0.06)	0.79 (0.17)	0.01 (0.17)	
3 (0.86-1)	0.98 (0.04)	0.86 (0.19)	-0.12 (0.19)	

Figure 7-1 Age distribution of responders and non-responders

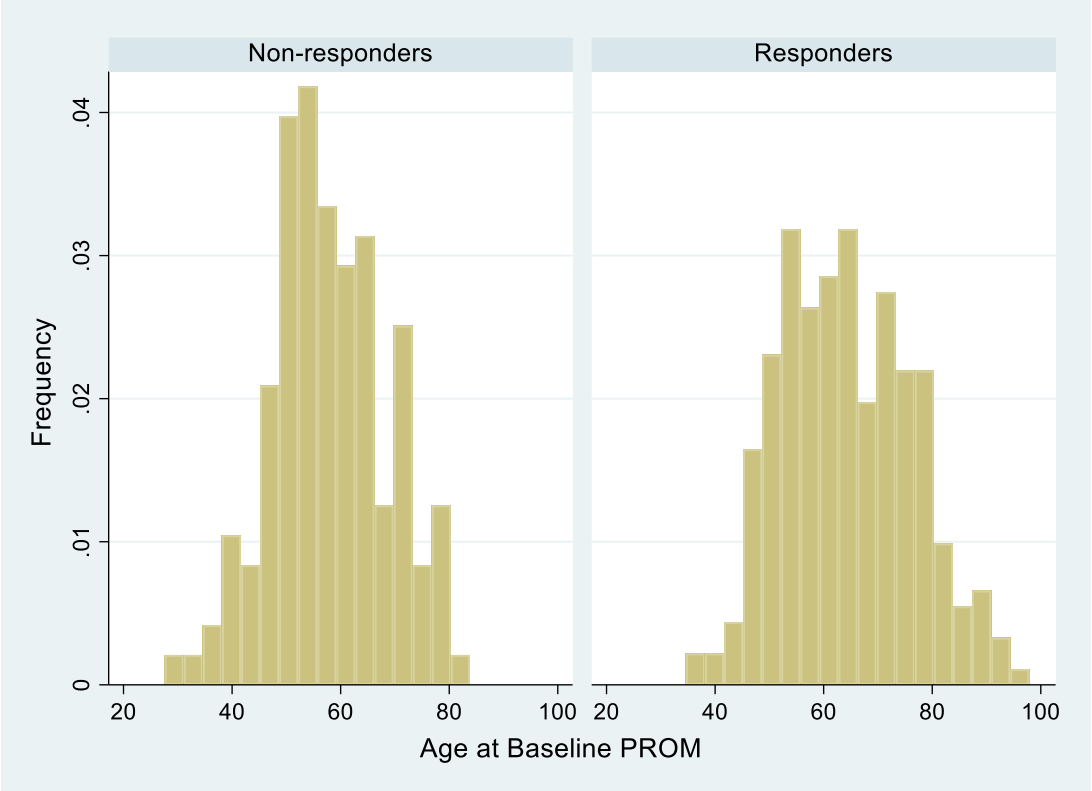
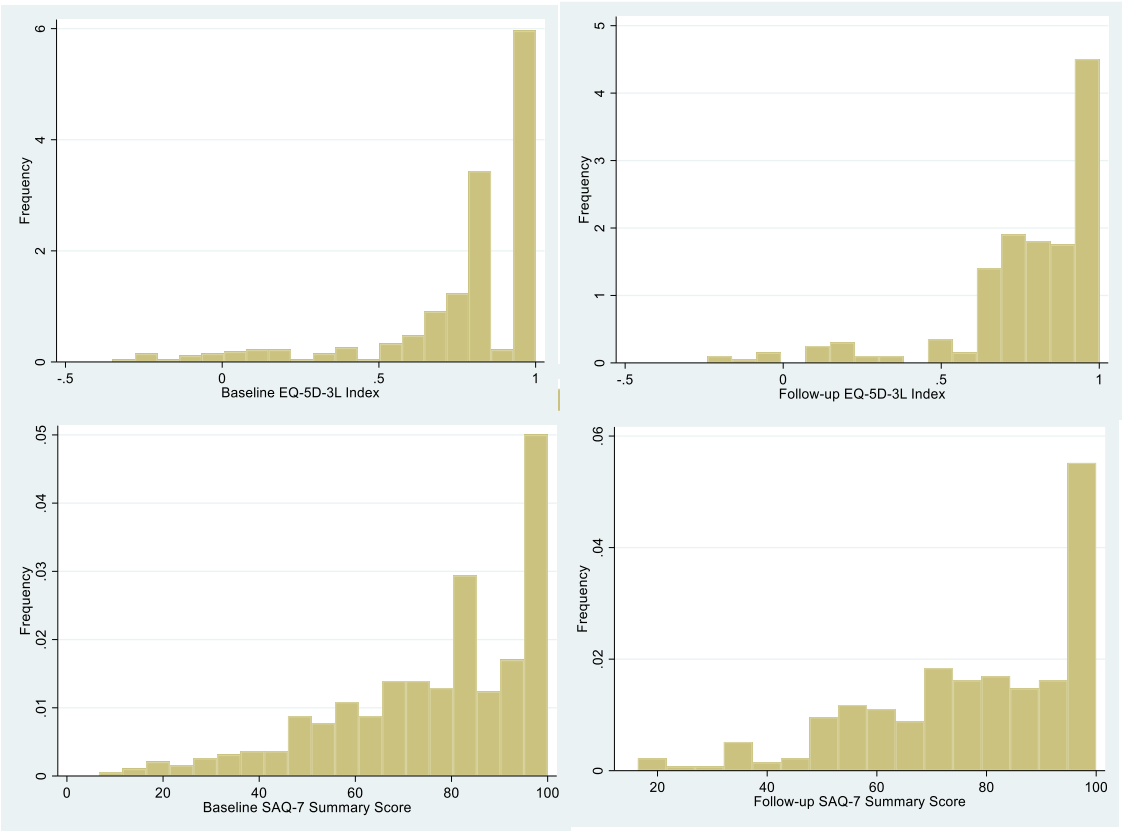


Figure 7-2 SAQ-7 summary and EQ-5D index score distributions



References

1 NHS England. A&E Attendances and Emergency Admissions. May 2016 Monthly Report. https://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2016/06/Monthly-AE-Report-May-16.pdf. 2016.

2 Holt P, Sinha S, Ozdemir B, *et al.* Variations and inter-relationship in outcome from emergency admissions in England: a retrospective analysis of Hospital Episode Statistics from 2005-2010. *BMC Serv Res* 2014;**14: 270.**

3 Steventon A, Deeny S, Friebel R, *et al.* The Health Foundation. Emergency hospital admissions in England: which may be avoidable and how? 2018.

4 Department of Health. Emergency admissions to hospital: managing the demand. 2013.

5 Greenhalgh J, Dalkin S, Gooding K, *et al.* Functionality and feedback: a realist synthesis of the collation, interpretation and utilisation of patient-reported outcome measures data to improve patient care. *Heal Serv Deliv Res* 2017;**5**.

6 Black N. Patient reported outcome measures could help transform healthcare. *BMJ* 2013;**346**:19–21.

7 Tsevat J, Dawson N V, Wu AW, *et al.* Health values of hospitalized patients 80 years or older. HELP Investigators. Hospitalized Elderly Longitudinal Project. *JAMA* 1998;**279**:371–5.

8 Chung S, Gale CP, Sundström J. Comparison of hospital variation in acute myocardial infarction care and outcome between Sweden and United Kingdom: population based cohort study using nationwide clinical registries. *BMJ* 2015;**351**:h3913.

9 Rinfret S, Grines CL, Cosgrove RS, *et al.* Quality of life after balloon angioplasty or stenting for acute myocardial infarction. One-year results from the Stent-PAMI trial. *J Am Coll Cardiol* 2001;**38**:1614–21.

10 Arnold S V, Chan PS, Jones PG, *et al.* Translational Research Investigating Underlying disparities in acute Myocardial infarction Patients' Health status (TRIUMPH): Design and Rationale of a Prospective Multicenter Registry. *Circ Cardiovasc Qual Outcomes* 2011;**4**:467–76.

11 Spertus JA, Peterson E, Rumsfeld JS, *et al.* The Prospective Registry Evaluating Myocardial Infarction: Events and Recovery (PREMIER)--evaluating the impact of myocardial infarction on patient outcomes. *Am Heart J* 2006;**151**:589–97.

12 Dodson JA, Geda M, Krumholz HM, *et al.* Design and rationale of the comprehensive evaluation of risk factors in older patients with AMI (SILVER-AMI) study. *BMC Health Serv Res* 2014;**14**:506.

13 Kwong E, Black N. Feasibility of collecting retrospective patient reported outcome measures (PROMs) in emergency hospital admissions. *Manuscr Submitt J Patient Rep Outcomes* 2018.

14 Kwong E, Black N. Retrospectively patient-reported pre-event health status showed strong association and agreement with contemporaneous reports. *J Clin Epidemiol* 2017;**81**:22–32.

15 Kwong E, Neuburger J, Black N. Agreement between retrospectively and contemporaneously collected patient-reported outcome measures (PROMs) in hip and knee replacement patients. *Qual Life Res* Published Online First: February 2018.

16 Runkel PJ, McGrath JE. *Research on Human Behavior: A Systematic Guide to Method.* American Educational Research Association 1973.

17 Chan PS, Jones PG, Arnold SA, *et al.* Development and validation of a short version of the Seattle angina questionnaire. *Circ Cardiovasc Qual Outcomes* 2014;**7**:640–7.

18 Spertus JA, Winder JA, Dewhurst TA, *et al.* Development and evaluation of the Seattle Angina questionnaire: a new functional status measure for coronary artery disease. *J Am Coll Cardiol* 1995;**25**:333–41.

19 Dougherty CM, Dewhurst T, Nichol WP, *et al.* Comparison of three quality of life instruments in stable angina pectoris: Seattle Angina Health Survey (SF-36), and Quality of Life Index. *Card Epidemiol* 1998;**51**:569–75.

20 McNamara RL, Spatz ES, Kelley TA, *et al.* Standardized Outcome Measurement for Patients With Coronary Artery Disease: Consensus From the International Consortium for Health Outcomes Measurement (ICHOM). *J Am Hear Assoc* 2015;**4**.

21 EuroQol Group. EuroQol - a new facility for the measurement of health-related quality of life. *Health Policy (New York)* 1990;**16**:199–208.

22 Dolan P, Roberts J. Modelling valuations for Eq-5d health states: an alternative model using differences in valuations. *Med Care* 2002;**40**:442–6.

23 Department for Communities and Local Government. English indices of deprivation. Crown Copyr. 2015. 2015.

24 Spertus JA, McDonell M, Woodman CL, *et al.* Association between depression and worse disease-specific functional status in outpatients with coronary artery disease. *Am Heart J* 2000;**140**:105–10.

25 Grodzinsky A, Arnold S V, Gosch K, *et al.* Angina frequency after acute myocardial infarction in patients without obstructive coronary artery disease. *Eur Hear J - Qual Care Clin Outcomes* 2015;**1**:92–9.

26 Roger VL. Epidemiology of heart failure. *Circ Res* 2013;**113**:646–59.

27 Cahill TJ, Kharbanda RK. Heart failure after myocardial infarction in the era of primary percutaneous coronary intervention: Mechanisms, incidence and identification of patients at risk. *World J Cardiol* 2017;**9**:407–15.

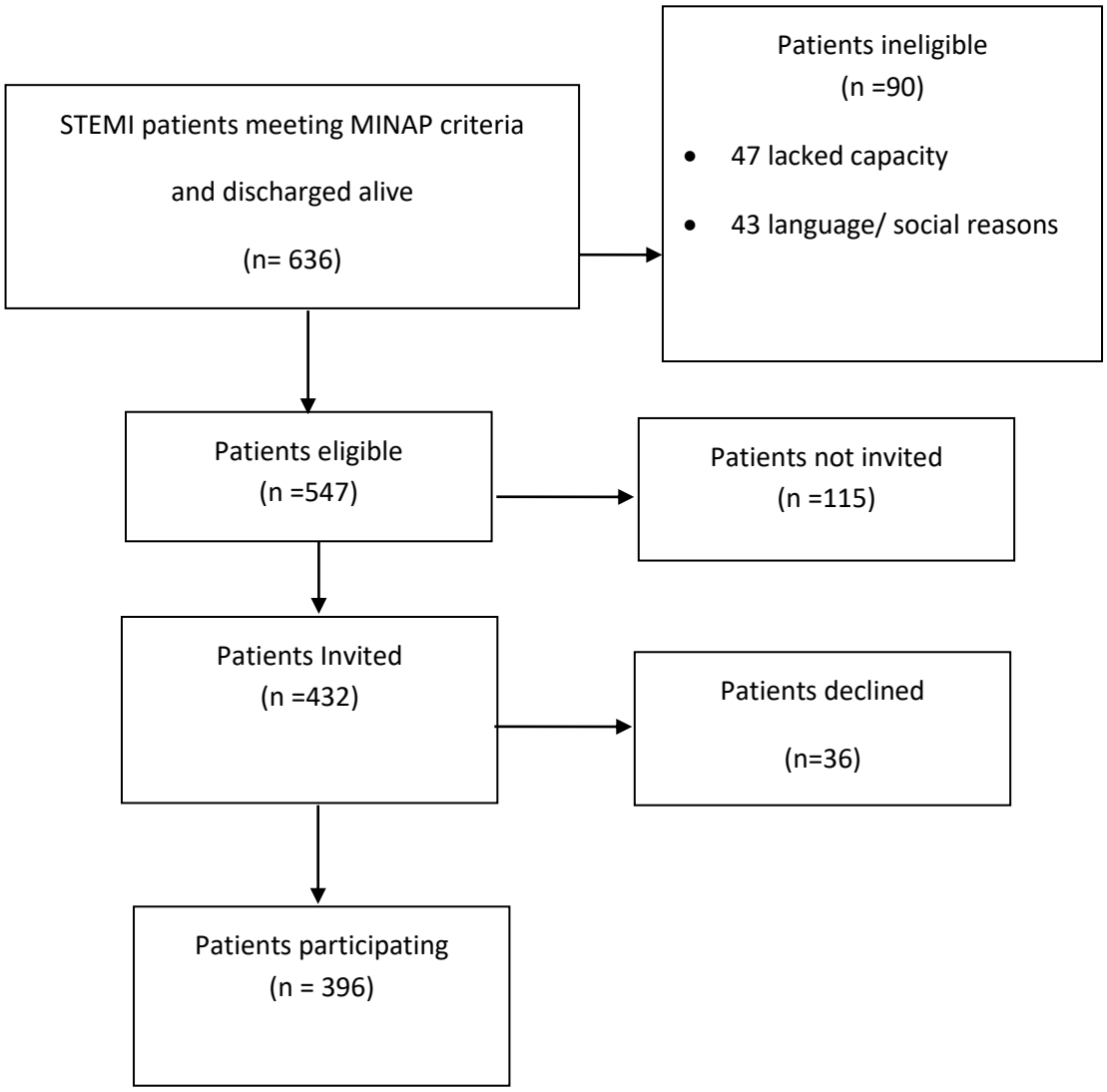
28 Sprangers M. Response-shift bias: a challenge to the assessment of patients' quality of life in cancer clinical trials. *Cancer Treat Rev* 1996;**22 Suppl A**:55–62.

29 Williams RB. Depression After Heart Attack. *Circulation* 2011;**123**:e639–40.

30 Norman G, Streiner D. *PDQ Statistics*. 3rd ed. BC Becker Inc., Hamilton, London. 2003.

31 Minitab Blog Editor. Choosing Between a Nonparametric Test and a Parametric Test. [ONLINE]. Minitab Blog Ed. 2015.

Appendix: Study Flow Diagram



Chapter 8: Discussion

8.1 Introduction

This final chapter of my thesis summarises the contribution of this work to our knowledge about patient reported outcome measures, and the feasibility of using PROMs to evaluate emergency health care.

Given the need to further our understanding of outcomes after emergency admissions, and despite the various theoretical challenges (Chapter 1), I identified reliable methods that could be used to establish patients' baseline health status when contemporary PROMs cannot be collected; first by comparing the agreement between retrospective PROMs and contemporary PROMs and then by exploring use of PROMs from matched groups of respondents to the GP Patient Survey. Finally, I tested the feasibility of collecting PROMs in emergency admissions in a range of acute NHS hospitals in England to establish its acceptability and interpretability for use in routine patient care. I also set out the recommended next steps for this research, and discuss the policy implications for the assessment of health service effectiveness of patient care.

8.2 Key findings

The first objective of the thesis was to review the literature and summarise existing evidence on using retrospectively-collected PROMs or data from population-based surveys to determine baseline health status in patient cohorts. This review covered studies assessing agreement between retrospective and contemporary PROMs, with agreement quantified using intra-class correlation coefficients (ICCs) for continuous measures, and kappa coefficients for categorical measures. This published review article contributed to the scientific body of knowledge with several key findings. Overall, the evidence suggested that there is strong agreement between PROMs collected retrospectively and contemporaneously. Agreement was higher for continuous compared to categorical measures (ICCs > 0.75, kappas ranging from 0.3 to 0.6), for indices rather than individual items, and for retrospective PROMs with

shorter versus longer recall time intervals. The directions of any differences between retrospective and contemporary PROM responses showed no consistent pattern. The review of the literature suggested that both retrospective PROMs and population values can have a role to play when contemporary PROMs are not possible (Chapter 2). From the literature, our knowledge of the use of retrospective PROMs is limited. This leads me to the second objective to further investigate the relationship between retrospective and contemporary PROMs to explore the influence of patient characteristics and contextual factors. In order for PROMs to be used for emergency admissions in the NHS, first we need to be able to compare and understand the reasons for any differences in context and patient characteristics in order to establish suitable uses of retrospective PROMs when contemporary PROMs capture is not possible.

These findings guided the development of my methods for comparing retrospective and contemporary PROMs in the English NHS, and led me to conduct my cohort study that addressed the second objective, to compare the agreement between retrospective and contemporary PROMs in elective conditions. The choice to focus on elective patient cohorts allowed me to make use of existing contemporary PROMs collected from patients before their surgery, by linking with the national PROMs programme data. I conducted the additional collection of retrospective PROMs to enable a direct comparison of the relationship between contemporary and retrospective PROMs. The study found strong agreement between retrospective and contemporary disease-specific PROMs and EQ-5D in elective orthopaedic patients collected in the English NHS, with ICCs of 0.8 for the disease-specific PROMs (OHS and OKS), and 0.6 for the EQ-5D. Although patients reported slightly lower scores in the retrospective questionnaire compared to the contemporary, the differences were small and none were statistically significant. I found that the strength of agreement was consistently high, regardless of the severity of a patient's condition, and social characteristics (age and SES) had a small effect (agreement was slightly lower in over-75s) or no significant influence. Mean retrospective PROMs for groups of patients could also reliably predict their mean contemporary PROM scores (Chapter 3). These study findings

confirmed the findings in the literature, supporting the potential use of retrospective PROMs in patient cohorts. In addition, the experience and knowledge of designing a study protocol and conducting a study to successfully administer a retrospective PROM questionnaire for elective patients during their inpatient hospital admission then informed the subsequent development of the design and protocol for the feasibility studies of collecting retrospective PROMs in emergency patients.

The third objective was to determine whether population-based values from the GP Patient Survey (GPPS) could be used to form baseline PROMs for patient cohorts, as a cheaper alternative to retrospective PROMs data collection. This study explored methods for matching contemporary and retrospective PROMs in elective patient cohorts to population groups of GPPS respondents with similar characteristics. It compared mean EQ-5D index scores across the different groups. Although differences between patients' contemporary and retrospective EQ-5D scores were small, these scores were very different to the mean EQ-5D scores of matched GPPS populations. These differences persisted after accounting for a range of patient characteristics matched to the population sample. I also found that restriction of matching by local authority did not narrow the differences when compared with matches using national data. The use of the latter would be preferable as a larger sample is achievable and the place of residence of patients is not required. Specific matching of co-morbidities to exact conditions also did not narrow the differences in mean EQ-5D scores compared to a count of the number of co-morbidities. The exploratory methods from this section of the thesis provided insights into the potential of using population values by matching to patient characteristics, made possible because these characteristics are routinely collected by the annual GPPS (Chapter 4). Although the findings presented in this chapter are not supportive of using population-based PROMs to form baseline scores for elective patient cohorts. However, there are particular reasons in elective patients such as long-standing conditions that warrant surgery; in this case arthritis of the hip or knee which may have lowered the baseline health statuses of these patients, in comparison to the general population. Therefore, there may be value to further exploration of these

comparisons in emergency patient cohorts, particularly those who experienced sudden illness or injury.

The fourth objective was to test the feasibility of collecting retrospective PROMs in emergency admissions, conducted in two separate patient cohorts, emergency laparotomy (EL) and STEMI. The findings from these studies provided an insight into the acceptability of recruiting patients to complete PROMs in emergency admissions. I identified factors that staff encountered when recruiting patients during their emergency admissions at each stage of identifying eligible patients, inviting them and patients' participation. This enabled recommendations to optimise these processes to be made, in view of future uses of PROMs in these contexts. This knowledge can therefore support routine practice of PROMs collections and further research in the NHS to improve the organisation in the collection of PROMs in emergency admissions to maximise recruitment rates (Chapter 5).

Furthermore, the findings regarding follow-up PROMs three months later for emergency patients demonstrated that this was acceptable in such patients with response rates of 74% (EL) and 66% (STEMI). On average patients regained their prior level of health status when measured using a disease-specific PROM. Differences in change scores by patient characteristics were slight, suggesting minimal response bias. These studies provided insight into the interpretability of PROMs in these contexts (Chapter 6 & 7).

8.3 Limitations

First, reliability between contemporary and retrospective PROMs was tested in elective patient cohorts, rather than patients undergoing emergency surgery. This was unavoidable as recognised in the theoretical assumptions of the thesis (Chapter 1). It is not possible to test reliability of contemporary PROMs and retrospective PROMs directly in emergency admissions since it is not possible to collect the former. To make use of findings from this first part of the PhD study, one must assume that the findings in the reliability of retrospective PROMs as tested in elective admissions holds true for emergency admissions.

Some support for this assumption can be drawn from the feasibility study which demonstrated plausible health status recovery following EL and STEMI.

A second limitation is that although this thesis has shown that it is feasible to collect retrospective PROMs in two emergency hospital admission cohorts, it cannot be assumed that this is equally feasible in other emergency patient cohorts. There remains the need to test for feasibility in other emergency patients to establish recruitment and response rates.

A third limitation is the generalisability of those recruited for PROMs. In the case of emergency admissions, there may be a subset of patients admitted with frailty and cognitive impairments for whom PROMs may not be feasible without further appropriate arrangements, such as interviewer administered PROMs or by proxy. It is also the case that these more elderly patients (over 85 years old, frail, multi-morbidity, and cognitive impairment) are known to be an increasing proportion of those admitted as emergencies [1,2].

A fourth limitation is that collecting PROMs in emergency admissions will inevitability be limited by the availability of validated PROMs for the specific condition or intervention. Although the use of generic PROMs such as the EQ-5D, allows for comparisons between different patient cohorts, it is less sensitive to improvements compared with a disease-specific PROM as shown in Chapter 6 and 7. In addition, agreement between contemporary and retrospective PROMs is consistently stronger in disease-specific PROMs (Chapter 2 and 3). However, these PROM tools will need to be validated to measure and capture the acute health changes.

In both emergency PCI and EL cohorts, there were few PROMs that have been validated for these patients. Although the questionnaires used in the feasibility studies were deemed the most suitable option after reviewing the literature and discussions with clinical stakeholders, both had limitations. The SAQ-7, although previously validated for use in AMI, was not specifically developed for AMI and can be used for other conditions such as stable and unstable angina. Similarly, there were limitations with the GIQLI, which was chosen as it was the most commonly used validated tool in emergency

abdominal surgery and system specific. It was the most suitable questionnaire for EL with patients who were admitted with a range of gastrointestinal conditions. A limitation I found during the study was the lack of the 'not applicable' option for the sexual health item. Another example is that for some items the word 'anus' was used, this was not appropriate for some patients at the follow-up questionnaire if their EL led to a colostomy or ileostomy.

Furthermore, these PROMs were not specifically develop for emergency admissions and retrospective use e.g. the GIQLI uses wording about 'your illness' in some items which may not be relevant and thus confusing for some EL patients presenting acutely with no known prior 'illness'. PROMs suitable for use retrospectively in emergency admissions will need to be developed or current questionnaires modified, with sufficient re-validation to enable routine use.

Another possibility is the development of new PROMs specifically for emergency patients which does not assume pre-existing steady symptom states. The ideal PROM will also have a suitable recall period and be flexible with options for the use of past tense wording so it can be applied retrospectively for baseline measurement.

8.4 Implications for policy and practice

8.4.1 Expanding beyond process measures and mortality outcomes

Emergency admissions were once seen primarily as lifesaving interventions, however, since survival following emergency acute intervention and hospital care is increasing, quality of life is becoming an increasingly crucial and relevant outcome.

Measuring PROMs can provide a unique opportunity for greater insight into quality of life outcomes following emergency hospital admission and treatment, further understanding the variation between providers, and complement the clinical measures that clinical audits currently collect [3,4]. The systematic development of PROMs in clinical areas where there is increasing demand, where variation in quality is still relatively unknown is paramount [5,6], so that

effectiveness and equity can be optimally addressed. In this thesis, I addressed the challenges in methods and feasibility that have resulted in the current lack of routine use of PROMs in emergency admissions. I have shown that a reliable and feasible method for doing so is achievable within the context of our socially funded NHS health system.

Next steps for the NHS should be promoting research to establish whether there is variation in quality of life following emergency admissions. Unwarranted variation may be a signal of disparities in the quality and effectiveness of healthcare between providers, and this would be valuable to identify. PROMs have a role as a quality indicator, allowing benchmarking in conjunction with other forms of quality data for the NHS.

8.4.2 Clinical benchmarking and quality indicator

Another purpose for collecting PROMs data is to enable local benchmarking, and expanding to provider comparisons to aid understanding of any unwarranted variation in outcomes. This will distil areas of attention for local providers of hospital care, as well as commissioners, and NHS England's overall national oversight in reducing unwarranted variation.

For patients admitted with emergencies, there is likely to be more variability in their acute presenting conditions compared with elective care; more processes will therefore be involved at the point of admission when urgent treatment is required. The aim of health services is to restore patients' health to their baseline status. Following emergency admissions, some patients will not be able to return to their baseline for several reasons. Firstly, it could be due to their underlying primary and co-morbid conditions. Secondly, the severity of their acute episode has an impact on recovery. Thirdly, the effectiveness of the treatment or care received is also relevant. In the context of emergency admissions, we need to use PROMs to allow comparisons based on the effectiveness of the service received. Therefore we would want to link PROMs to clinical audit data to better case-mix adjust and standardise accounting for the differences in outcomes due to the first two reasons.

The ways in which PROMs can be used in emergency admissions is as aggregate measures of service quality at the hospital, informing service level improvements and provider comparisons. There may be some inherent differences compared to elective admissions when used for these purposes. In both areas, the objective is to understand the effectiveness of treatment and care following hospital interventions.

In emergency care however, the patient journey is generally more complex, involving more coordination and transfer of care between different clinical teams and specialty areas both within the hospital (e.g. Emergency Department, Intensive Care Units), and outside of the hospital (e.g. pre-hospital care, ambulance services) and in many cases patients have a more convoluted post discharge and rehabilitation period.

If PROMs can be utilised and interpreted effectively, this can be a valuable instrument for the assessment of effectiveness of acute services across the patient pathway and therefore could reflect the effectiveness of the hospital organisation encompassing multiple pathways and departments than compared to PROMs in elective care. As such, bringing about pathway and service improvements will also require the involvement of more than a single team or department. PROMs for emergency care should therefore be disseminated in a way that reaches across departmental boundaries and professional silos; it will then fulfil its potential to serve as a broader quality indicator for across a range of hospital services.

8.4.3 PROMs for patient care

Deaths following acute myocardial infarction admissions have fallen and similar improvements have been seen in survival following emergency laparotomy since the start of the national project (NELA). National clinical audits have helped the understanding of these trends and associations [7,8].

Clinicians currently use audits to understand morbidity and mortality, but PROMs will add another dimension to their knowledge of longer-term health status (including HRQL) during their recovery after treatment, and if they regain their prior health. This additional knowledge can help support clinicians

to make continued improvements in holistic patient care and shared decision making regarding treatment options, through timely feedback of PROMs results embedded within clinical audit data. The feasibility studies (Chapters 6 and 7) have given us insight into the health status recovery at three months in those groups of patients.

Further uptake and benefits of PROMs could be enhanced through linking to national clinical audits. Linking and embedding PROMs for emergency admissions to national clinical audits enables clinical data to be used in case-mix adjustment models for interpreting PROMs scores. It can also facilitate understanding of the relationship between clinical parameters during the index hospitalisation with patients' long term recovery and quality of life. For example, PROMs data could be linked to clinical measures known to be associated with outcome (such as Kilip classification, concentration of Troponin I, infarct site in STEMI and P- POSSUM scores for EL).

Embedding PROMs in national audits and registries also has the benefit of engaging the clinical community to measure and use PROMs in a way that supports clinical management, increasing clinicians' familiarity and ownership in their role as valuable outcomes alongside morbidity, mortality and process measures in the clinical management of patients.

8.5 Implications for research

The findings of this thesis lead to the following further research topics for considering whether routine collection of PROMs in emergency admissions is justifiable for the NHS.

8.5.1 Exploring use of GPPS PROMs as a proxy baseline in emergency admission cohorts of patients

Based on the hypothesis that patients admitted with sudden emergency admissions may have a baseline health status similar to those in the general population, the next step is to conduct studies with retrospective PROMs in emergency patient cohorts (EL and STEMI patients), matched to the GPPS population, using methods already explored from this thesis (Chapter 4).

8.5.2 Describing the variation in outcomes following emergency admissions and comparing providers using PROMs

This would establish whether there is variation in PROMs following emergency admissions; establish methods to collect PROMs data in large samples and hospitals to allow provider comparisons, and develop approaches to embed and publish PROMs results in a way that allows for dissemination and benchmarking through national clinical audit.

The objectives would be to:

- Expand PROMs collection in a large-scale study with EL and STEMI patients, in more hospital sites and for longer periods. The studies should aim to recruit circa. 3000 patients in each condition and involve 30 hospitals. The study sample will be powered to enable comparisons across providers.
- To link patient level PROMs data to national clinical audit data to understand the relationship between patient admission characteristics and treatment process measures collected by national clinical audits. In order to explore and understand how PROMs can be used to inform clinical decision making in emergency admissions (e.g. for used in the development of use in risk prediction models).
- To investigate the relationship between survival following emergency admission and quality of life of survivors.

8.5.3 Establish best methods and modality to collect PROMs in emergency admissions.

This will evaluate the best methods and modality for the routine collection of PROMs in emergency admissions, by testing different modes of data collection methods such as electronic, and different follow-up time points. The objectives would be to:

- Test alternative modes of delivery alongside paper-based in EL and STEMI patients.
- Ascertain optimal time points for follow-up in EL and STEMI patients.

- Explore the best models of case-mix adjustment models in EL and STEMI patients.
- Establish the relative cost of different modalities in the routine collection of PROMs for emergencies in the NHS.

8.5.4 Expansion into other emergency conditions:

This fourth aim is to establish further types of emergency admissions for which PROMs collection is both useful and feasible as a quality indicator. This would include development and feasibility testing PROMs in other emergency conditions that are more atypical to STEMI and EL. e.g. in stroke patients.

8.6 References

1 National Audit Office. Reducing Emergency Admissions Report to Dept of Health & Social Care NHS England. 2011.

2 Steventon A, Deeny S, Friebel R, *et al.* The Health Foundation. Emergency hospital admissions in England: which may be avoidable and how? 2018.

3 Greenhalgh J, Dalkin S, Gooding K, *et al.* Functionality and feedback: a realist synthesis of the collation, interpretation and utilisation of patient-reported outcome measures data to improve patient care. *Heal Serv Deliv Res* 2017;**5**.

4 Devlin NJ, Appleby J. King's Fund. Getting the most out of PROMs: Putting health outcomes at the heart of NHS decision-making. 2010. http://www.kingsfund.org.uk/publications/proms.html

5 Browne JP, Cano SJ, Smith S. Using Patient-reported Outcome Measures to Improve Health Care. *Med Care* 2017;**55**:901–4.

6 Black N. Patient reported outcome measures could help transform healthcare. *BMJ* 2013;**346**:19–21.

7 NELA Project Team. Third Patient Report of the National The third Patient Report of the National Emergency Laparotomy Audit (NELA). 2017.

8 Herrett E, Smeeth L, Walker, Lynne *et al.* Myocardial Ischaemia National Audit Project. National Institute for Clinical Outcomes Research 2016.

Thesis Appendices

9.1 Retrospective PROMs Questionnaire for Hip Arthroplasty

Study ID ☐☐-☐☐☐-☐-☐

Hip Surgery Questionnaire

Retrospective Questionnaire

PATIENT BOOKLET

Project team:

Dr Esther Kwong (Chief Investigator and project lead)
Professor Nick Black (PhD Supervisor)

London School of Hygiene and Tropical Medicine (LSHTM)

Local hospital project leads:

[Insert names of local research team]

Study ID ☐☐-☐☐☐-☐-☐

Participant Consent Form

Please read the information,
Initial the relevant boxes and sign overleaf.

Please Sign your Initials in boxes

I have read and understood the enclosed Participant Information Sheet [INSERT(version X.X date)]. I have had the opportunity to consider the information and have had my questions answered.	☐
I understand that if I complete this questionnaire I will not be identified by name in any published reports or papers.	☐
I understand that I am free to withdraw from taking part at any time, without giving a reason.	☐
I understand that all information I provide will be kept confidentially.	☐
I understand that sections of my medical notes may be looked at by responsible individuals as part of the study and information may be passed to the research team. I give permission for these individuals to have access to my records. I also understand that sections of my medical records may be looked at as part of study monitoring, or from the ethics committee, or regulatory authorities where it is relevant to my taking part in this research.	☐
I understand that my personal details will be held and stored by the research team at London School of Hygiene & Tropical Medicine and may be used by the evaluation study to link the questionnaire to information that is routinely collected on the national PROMs programme.	☐
I understand that my questionnaire will be given a unique number that will also be on my consent form, and may be used to link my data.	☐
I understand that the research team will not release my personal details, unless required by law. In such an exceptional event I will be told if any disclosure will take place.	☐

If you have agreed to take part in this study please write your name and address in CAPITAL LETTERS below.

Title _____

First Name _____

Surname _____

Address _____

_____

_____

Postcode _____

Date of birth ___/__/____

Signature _____

Name (in capital letters) _____

Today's date ___/__/____

FOR COMPLETION BY STAFF:

Affix Patient Label with NHS number in this space

Please ensure patient label contains *their NHS number*

Name of staff taking consent (in capital letters)

_____

Signature _____

Today's date ___/__/____

Please now tear out this page

Version 1.3: 19/02/16

Study ID ☐☐-☐☐☐-☐-☐

Retrospective Patient Reported Outcomes
Hip Surgery Questionnaire

Q1. What is your date of birth?

D	D	M	M	Y	Y	Y	Y

Please ensure this is your **date of birth NOT** today's date

Q2. Is anyone helping you fill in this questionnaire?

Yes ☐ 1 No ☐ 2

If the answer is yes, please give the relationship to you of the person assisting you

Q3. Are you?

Male ☐ 1 Female ☐ 2

Q4. Which statement best describes your living arrangements?

I live with partner/ spouse/ family/ friends ☐ 1

I live alone ☐ 2

I live in a nursing home, hospital or other long-term care home ☐ 3

Other ☐ 4

Q5 For how long have you had problems with the hip on which you have had surgery?

Less than1 year	1 to 5 years	6 to 10 years	More than 10 years
☐ 1	☐ 2	☐ 3	☐ 4

Q6 Have you had a previous joint replacement on the hip on which you have just had surgery on?

Yes ☐ 1 No ☐ 2

Hip Surgery Questionnaire –

We are interested in finding out about the problems you have had with the hip on which you have had surgery. Please let us know how you were before your operation

Tick one box for every question.

Q7. During the past 4 weeks before your operation...

How would you describe the pain you usually had from your hip?

None	Very mild	Mild	Moderate	Severe
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q8. During the past 4 weeks before your operation...

Did you have any trouble with washing and drying yourself (all over) because of your hip?

No trouble at all	Very little trouble	Moderate trouble	Extreme difficulty	Impossible to do
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q9. During the past 4 weeks before your operation...

Did you have any trouble getting in or out of a car or using public transport because of your hip? (Whichever you tend to use)

No trouble at all	Very little trouble	Moderate trouble	Extreme difficulty	Impossible to do
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q10. During the past 4 weeks before your operation...

Have you been able to put on a pair of socks, stockings or tights?

Yes easily	With little difficulty	With moderate difficulty	With extreme difficulty	No, impossible
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q11. During the past 4 weeks before your operation...

Could you do the household shopping <u>on your own</u>?

Yes easily	With little difficulty	With moderate difficulty	With extreme difficulty	No, impossible
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q12. During the past 4 weeks before your operation...

For how long had you been able to walk before <u>pain from your hip</u> becomes severe? (With or without a stick)

No pain/More than 30 minutes	16 to 30 minutes	5 to 15 minutes	Around the house *only*	Not at all – pain severe on walking
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q13. During the past 4 weeks before your operation...

Have you been able to climb a flight of stairs?

Yes easily	With little difficulty	With moderate difficulty	With extreme difficulty	No, impossible
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q14. During the past 4 weeks before your operation...

After a meal (sat at a table), how painful had it been for you to stand up from a chair <u>because of your hip</u>?

Not at all painful	Slightly painful	Moderately painful	Very painful	Unbearable
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q15. During the past 4 weeks before your operation...

Had you been limping when walking, <u>because of your hip?</u>

Rarely/ Never	Sometimes or just at first	Often, not just at first	Most of the time	All of the time
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q16. During the past 4 weeks before your operation...

Have you had any sudden severe pain –'shooting', 'stabbing', 'spasms'- from the affected hip?

No days	Only 1 or 2 days	Some days	Most days	Every day
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q17 During the past 4 weeks before your operation...

How much had pain from your hip interfered with your usual work (including housework)?

Not at all	A little bit	Moderately	Greatly	Totally
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q18. During the past 4 weeks before your operation...

Had you been troubled by pain from your hip in bed at night?

No nights	Only 1 or 2 nights	Some nights	Most nights	Every night
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

By placing a tick in one box in each group below, please indicate which statements best describe your own health before your operation.

19. Mobility

I had no problems in walking about	☐ 1
I had some problems in walking about	☐ 2
I was confined to bed	☐ 3

20. Self-Care

I had no problems with self-care	☐ 1
I had some problems washing or dressing myself	☐ 2
I was unable to wash or dress myself	☐ 3

21. Usual Activities *(e.g. work, study, housework, family or leisure activities)*

I had no problems with performing my usual activities	☐ 1
I had some problems with performing my usual activities	☐ 2
I was unable to perform my usual activities	☐ 3

22. Pain/Discomfort

I had no pain or discomfort	☐ 1
I had moderate pain or discomfort	☐ 2
I had extreme pain or discomfort	☐ 3

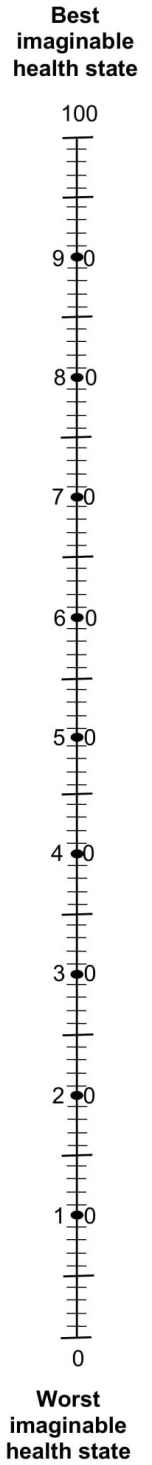
23. Anxiety/Depression

I was not anxious or depressed	☐ 1
I was moderately anxious or depressed	☐ 2
I was extremely anxious or depressed	☐ 3

24.
To help people say how good or bad a health state is, we have drawn a scale (rather like a thermometer) on which the best state you can imagine is marked 100 and the worst state you can imagine is marked 0.

We would like you to indicate on this scale how good or bad your own health was before your operation, in your opinion. Please do this by drawing a line from the box below to whichever point on the scale indicates how good or bad your health state <u>was before your operation</u>

Your own health state before your operation



25.How long has it been since you had surgery to your hip?

One day ☐1
Two days ☐2
Three days ☐3
Four days ☐4
Five days ☐5
Other _____

26.Today's Date

D D M M 2 0 Y Y

Please ensure this **is today's date**
NOT your date of birth

You have now finished the questionnaire.

Thank you for your help.
Please return this booklet to [insert local project team names] or the clinical staff that invited you to this study

9.2 Retrospective PROMs Questionnaire for Knee Arthroplasty

Study ID ☐☐-☐☐☐-☐-☐

Knee Surgery Questionnaire

Retrospective Questionnaire

PATIENT BOOKLET

Project team:

Dr Esther Kwong (Chief Investigator and project lead)
Professor Nick Black (PhD Supervisor)

London School of Hygiene and Tropical Medicine (LSHTM)

Local hospital project leads:

[Insert names of local research team]

Study ID ☐☐-☐☐☐-☐-☐

Participant Consent Form

Please read the information,
Initial the relevant boxes and sign overleaf.

Please Sign your Initials in boxes

I have read and understood the enclosed Participant Information Sheet [INSERT(version X.X date)]. I have had the opportunity to consider the information and have had my questions answered.	☐
I understand that if I complete this questionnaire I will not be identified by name in any published reports or papers.	☐
I understand that I am free to withdraw from taking part at any time, without giving a reason.	☐
I understand that all information I provide will be kept confidentially.	☐
I understand that sections of my medical notes may be looked at by responsible individuals as part of the study and information may be passed to the research team. I give permission for these individuals to have access to my records. I also understand that sections of my medical records may be looked at as part of study monitoring, or from the ethics committee, or regulatory authorities where it is relevant to my taking part in this research.	☐
I understand that my personal details will be held and stored by the research team at London School of Hygiene & Tropical Medicine and may be used by the evaluation study to link the questionnaire to information that is routinely collected on the national PROMs programme.	☐
I understand that my questionnaire will be given a unique number that will also be on my consent form, and may be used to link my data.	☐
I understand that the research team will not release my personal details, unless required by law. In such an exceptional event I will be told if any disclosure will take place.	☐

If you have agreed to take part in this study please write your name and address in CAPITAL LETTERS below.

Title _____

First Name _____

Surname _____

Address _____

_____

_____

Postcode _____

Date of birth ___/__/____

Signature _____

Name (in capital letters) _____

Today's date ___/__/____

FOR COMPLETION BY STAFF:

Affix Patient Label with NHS number in this space

Please ensure patient label contains *their NHS number*

Name of staff taking consent (in capital letters)

_____

Signature _____

Today's date ___/__/____

Please now tear out this page

Version 1.3: 19/02/16

Study ID ☐☐-☐☐☐-☐-☐

Retrospective Patient Reported Outcomes

Knee Surgery Questionnaire

Q1. What is your date of birth?

D	D	M	M	Y	Y	Y	Y

Please ensure this is your **date of birth NOT** today's date

Q2. Is anyone helping you fill in this questionnaire?

Yes ☐ 1 No ☐ 2

If the answer is yes, please give the relationship to you of the person assisting you

Q3. Are you?

Male ☐ 1 Female ☐ 2

Q4. Which statement best describes your living arrangements?

I live with partner/ spouse/ family/ friends ☐ 1

I live alone ☐ 2

I live in a nursing home, hospital or other long-term care home ☐ 3

Other ☐ 4

Q5 For how long have you had problems with the knee on which you have had surgery?

Less than 1 year	1 to 5 years	6 to 10 years	More than 10 years
☐ 1	☐ 2	☐ 3	☐ 4

Q6 Have you had a previous joint replacement on the knee on which you have just had surgery on?

Yes ☐ 1 No ☐ 2

Knee Surgery Questionnaire –

We are interested in finding out about the problems you have had with the knee on which you have had surgery. Please let us know how you were before your operation

Tick one box for every question.

Q7. During the past 4 weeks before your operation...

How would you describe the pain you usually had from your knee?

None	Very mild	Mild	Moderate	Severe
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q8. During the past 4 weeks before your operation...

Did you have any trouble with washing and drying yourself (all over) because of your knee?

No trouble at all	Very little trouble	Moderate trouble	Extreme difficulty	Impossible to do
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q9. During the past 4 weeks before your operation...

Did you have any trouble getting in or out of a car or using public transport because of your knee? (Whichever you tend to use)

No trouble at all	Very little trouble	Moderate trouble	Extreme difficulty	Impossible to do
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q10. During the past 4 weeks before your operation...

For how long had you been able to walk before <u>pain from your knee</u> becomes severe? (With or without a stick)

No pain/More than 30 minutes	16 to 30 minutes	5 to 15 minutes	Around the house *only*	Not at all – pain severe on walking
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q11. During the past 4 weeks before your operation...

After a meal (sat at a table), how painful had it been for you to stand up from a chair because of your knee?

Not at all painful	Slightly painful	Moderately painful	Very painful	Unbearable
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q12. During the past 4 weeks before your operation...

Had you been limping when walking, because of your knee?

Rarely/ Never	Sometimes or just at first	Often, not just at first	Most of the time	All of the time
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q13. During the past 4 weeks before your operation...

Could you kneel down and get up again afterwards?

Yes easily	With little difficulty	With moderate difficulty	With extreme difficulty	No, impossible
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q14. During the past 4 weeks before your operation...

Had you been troubled by pain from your knee in bed at night?

No nights	Only 1 or 2 nights	Some nights	Most nights	Every night
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q15. During the past 4 weeks before your operation...

How much had pain from your knee interfered with your usual work (including housework)?

Not at all	A little bit	Moderately	Greatly	Totally
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q16. During the past 4 weeks before your operation...

Had you felt that your knee might suddenly 'give way' or let you down?

Rarely/ Never	Sometimes or just at first	Often, not just at first	Most of the time	All of the time
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q17 During the past 4 weeks before your operation...

Could you do the household shopping on your own?

Yes easily	With little difficulty	With moderate difficulty	With extreme difficulty	No, impossible
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q18. During the past 4 weeks before your operation...

Could you walk down one flight of stairs?

Yes easily	With little difficulty	With moderate difficulty	With extreme difficulty	No, impossible
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

By placing a tick in one box in each group below, please indicate which statements best describe your own health before your operation.

19. Mobility

I had no problems in walking about	☐ 1
I had some problems in walking about	☐ 2
I was confined to bed	☐ 3

20. Self-Care

I had no problems with self-care	☐ 1
I had some problems washing or dressing myself	☐ 2
I was unable to wash or dress myself	☐ 3

21. Usual Activities *(e.g. work, study, housework, family or leisure activities)*

I had no problems with performing my usual activities	☐ 1
I had some problems with performing my usual activities	☐ 2
I was unable to perform my usual activities	☐ 3

22. Pain/Discomfort

I had no pain or discomfort	☐ 1
I had moderate pain or discomfort	☐ 2
I had extreme pain or discomfort	☐ 3

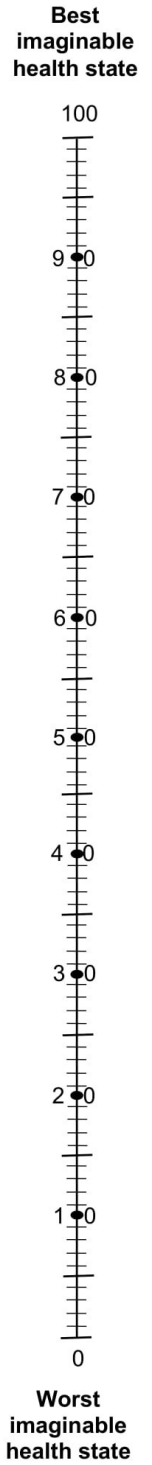
23. Anxiety/Depression

I was not anxious or depressed	☐ 1
I was moderately anxious or depressed	☐ 2
I was extremely anxious or depressed	☐ 3

24.
To help people say how good or bad a health state is, we have drawn a scale (rather like a thermometer) on which the best state you can imagine is marked 100 and the worst state you can imagine is marked 0.

We would like you to indicate on this scale how good or bad your own health was before your operation, in your opinion. Please do this by drawing a line from the box below to whichever point on the scale indicates how good or bad your health state <u>was before your operation</u>

Your own health state before your operation



25.How long has it been since you had surgery to your knee?

One day ☐
Two days ☐
Three days ☐
Four days ☐
Five days ☐
Other _____

26.Today's Date

D D M M 2 0 Y Y

Please ensure this **is today's date**
NOT your date of birth

You have now finished the questionnaire.

Thank you for your help.
Please return this booklet to [insert local project team names] or the clinical staff that invited you to this study

9.3 Retrospective PROMs Questionnaire for Emergency Laparotomy

Study ID ☐☐-☐☐☐-☐-☐

Version 1.4

Patients' views of the outcome of care

Emergency Laparotomy

Inpatient Questionnaire

Project team:

Dr Esther Kwong (Chief Investigator and project lead)
Professor Nick Black (PhD Supervisor)

London School of Hygiene and Tropical Medicine (LSHTM)

Local hospital project leads:

[Insert names of local research team]

Study ID ☐☐-☐☐☐-☐-☐

Consent Form

Please read the information, initial the relevant boxes and sign overleaf.

Please sign your initials in boxes

I have read and understood the enclosed Participant Information Sheet [INSERT(version X.X date)]. I have had the opportunity to consider the information and have had my questions answered.	☐
I understand that if participate in this study I will not be identified by name in any published reports or papers.	☐
I understand that I am free to withdraw from taking part at any time, without giving a reason.	☐
I understand that my personal details (address) will be used to contact me by post within 3-6 months to complete a follow-up questionnaire.	☐
I understand that sections of my medical notes may be looked at by responsible individuals as part of the study and information may be passed to the research team. I give permission for these individuals to have access to my records. I also understand that sections of my medical records may be looked at as part of study monitoring, or from the ethics committee, or regulatory authorities where it is relevant to my taking part in this research.	☐
I understand that my personal details (date of birth, postcode and NHS number) will be held and stored by the research team at London School of Hygiene & Tropical Medicine and will be supplied to NHS Digital (Previously the Health and Social care Information Centre) to check vital status before I am sent a follow-up questionnaire.	☐
I understand that my personal details (date of birth and NHS number) will be used by the study to link to information that is routinely collected by national clinical audits.	☐
I understand that my questionnaire will be given a unique number that will also be on my consent form, and may be used to link my data.	☐
I understand that the research team will not otherwise release my personal details, except for the above reasons, unless required by law. I will be told if any disclosure was to take place under an exceptional event.	☐

If you have agreed to take part in this study please write your name and address in CAPITAL LETTERS below.

Title _____

First Name _____

Surname _____

Address _____

_____

_____

Postcode _____

Date of birth ___/__/____

Signature _____

Today's date ___/__/2017

FOR COMPLETION BY STAFF:

Affix Patient Label with NHS number in this space

Please ensure patient label contains *their NHS number and Address*

Date of Admission ___/___ / 2017

Name of staff taking consent (in capital letters)

_____

Signature _____

Today's date ___/__/2017

Please now tear out this page

: Study ID ☐☐-☐☐☐-☐-☐

Patients' views of the outcome of care

Emergency Laparotomy

Inpatient Questionnaire

Q1. What is your date of birth?

D	D	M	M	Y	Y	Y	Y

Please ensure this is your **date of birth NOT** today's date

Q2. Is anyone helping you fill in this questionnaire?

Yes ☐ 1 No ☐ 2

If the answer is yes, please give the relationship to you of the person assisting you

Q3. Are you?

Male ☐ 1 Female ☐ 2

Q4. Which statement best describes your living arrangements?

I live with partner/ spouse/ family/ friends ☐ 1

I live alone ☐ 2

I live in a nursing home, hospital or other long-term care home ☐ 3

Other ☐ 4

Q5 For how long have you had problems with your abdomen (tummy) before your current admission?

Less than two weeks ☐ 1 Two weeks to One month ☐ 2 1 month to 12 months ☐ 3 More than 1 year ☐ 4

Q6 Have you had a previous problem with your abdomen (tummy) for which you were admitted to hospital?

Yes ☐ 1 No ☐ 2

We are interested in finding out how you were <u>one month before your current admission.</u> Please answer the questions according to how you were usually when you were at home before you came into hospital rather than how you are feeling now.

Tick <u>one</u> box for <u>every</u> question.

Q7. <u>One month before your admission,</u> how often did you have pain in the abdomen?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q8. <u>One month before your admission,</u> how often did you have a feeling of fullness in the abdomen?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q9. <u>One month before your admission,</u> how often did you have bloating (sensation of too much gas in the abdomen)?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q10. <u>One month before your admission,</u> how often did you have trouble by excessive passage of gas through the anus?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q11. One month before your admission, how often did you have trouble with strong burping or belching?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q12. One month before your admission, how often did you have trouble with gurgling noises from the abdomen?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q13. One month before your admission, how often did you have trouble with frequent bowel movements?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q14. One month before your admission, how often did you find eating to be a pleasure?

Never	Little of the time	Some of the time	Most of the time	All of the time
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q15. One month before your admission, to what extent did you restrict the kinds of food you ate?

Very much	Much	Somewhat	Little	Not at all
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q16. One month before your admission, how well were you able to cope with everyday stresses?

Extremely Poorly	Poorly	Moderately	Well	Extremely Well
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q17 One month before your admission, how often were you sad about being ill?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q18. One month before your admission, how often were you nervous or anxious about your illness?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q19. One month before your admission, how often were you happy with life in general?

Never	Little of the time	Some of the time	Most of the time	All of the time
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q20. One month before your admission, how often were you frustrated about your illness?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q21. One month before your admission how often were you tired or fatigued?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q22. One month before your admission, how often did you feel unwell?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q23. One month before your admission, over a week, did you wake up in the night?

Every night	5-6 nights	3-4 nights	1-2 nights	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q24. One month before your admission, were you troubled by any changes in your appearance?

A great deal	A moderate amount	Somewhat	A little bit	Not at all
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q25. One month before your admission, how much physical strength did you lose?

A great deal	A moderate amount	Somewhat	A little bit	Not at all
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q26. One month before your admission, to what extent did you lose your endurance?

A great deal	A moderate amount	Somewhat	A little bit	Not at all
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q27. One month before your admission, to what extent did you feel unfit?

Extremely Unfit	Unfit	Moderately Unfit	A little Unfit	Fit
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q28. One month before your admission, how often were you able to complete your normal daily activities (school, work, household)?

Never	Little of the time	Some of the time	Most of the time	All of the time
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q29. One month before your admission, how often were you able to take part in your usual patterns of leisure or recreational activities?

Never	Little of the time	Some of the time	Most of the time	All of the time
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q30. One month before your admission, how much were you troubled by the medical treatment of your illness?

Very much	Much	Somewhat	Little	Not at all
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q31. To what extent did your personal relations with people close to you (family or friends) worsened one month before your admission because of your illness?

Very much	Much	Somewhat	Little	Not at all
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q32. To what extent was your sexual life been impaired (harmed) one month before your admission because of your illness?

Very much	Much	Somewhat	Little	Not at all
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q33. One month before your admission, did you have trouble with fluid or food coming up into your mouth (regurgitation)?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q34. One month before your admission, how often did you feel uncomfortable because of your slow speed of eating?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q35. One month before your admission, how often did you have trouble swallowing your food?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q36. One month before your admission how often did you have trouble with urgent bowel movements?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q37. One month before your admission, how often did you have trouble with diarrhoea?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q38. One month before your admission, how often did you have trouble with constipation?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q39. One month before your admission, how often did you have you trouble with nausea?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q40. One month before your admission, how often did you have trouble with blood in the stool?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q41. One month before your admission, how often did you have trouble with heartburn?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q42. One month before your admission, how often did you have trouble with uncontrolled stools?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

We would like you to think back to <u>four weeks ago</u>, by placing a tick in one box in each group below, please indicate which statements best describe your own health <u>at that time.</u>

43. Mobility

I have no problems in walking about	☐ 1
I have some problems in walking about	☐ 2
I am confined to bed	☐ 3

44. Self-Care

I have no problems with self-care	☐ 1
I have some problems washing or dressing myself	☐ 2
I am unable to wash or dress myself	☐ 3

45. Usual Activities *(e.g. work, study, housework, family or leisure activities)*

I have no problems with performing my usual activities	☐ 1
I have some problems with performing my usual activities	☐ 2
I am unable to perform my usual activities	☐ 3

46. Pain/Discomfort

I have no pain or discomfort	☐ 1
I have moderate pain or discomfort	☐ 2
I have extreme pain or discomfort	☐ 3

47. Anxiety/Depression

I am not anxious or depressed	☐ 1
I am moderately anxious or depressed	☐ 2
I am extremely anxious or depressed	☐ 3

Study ID ☐☐-☐☐☐-☐-☐

48. Today's Date

D D M M 2 0 1 7

Please ensure this **is today's date**
NOT your date of birth

50. Have you been told by a doctor that you have any of the following? (Tick all that apply)

Heart disease (for example angina heart attack or heart failure)	☐1	Kidney disease	☐7
High blood pressure	☐2	Diseases of the nervous system (for example Parkinson's disease or multiple sclerosis)	☐8
Problems caused by stroke	☐3	Liver disease	☐9
Leg pain when walking due to poor circulation	☐4	Cancer (within the last 5 years)	☐10
Lung disease (for example asthma, chronic bronchitis or emphysema)	☐5	Depression	☐11
Diabetes	☐6	Arthritis	☐12

You have now finished the questionnaire.

Thank you for your help.

Please return this booklet to [insert local project team names] or the clinical staff that invited you to this study

If you have found this questionnaire, please post this back to:

[Insert local team address or chief investigator address]

9.4 Retrospective PROMs Questionnaire for STEMI

Study ID ☐☐-☐☐☐-2-☐

Version 1.4

Patients' views of the outcome of care

Cardiology

Inpatient Questionnaire

Project team:

Dr Esther Kwong (Chief Investigator and project lead)
Professor Sir Nick Black (PhD Supervisor)

London School of Hygiene and Tropical Medicine (LSHTM)

Hospital project lead:

[Insert names of local research teams]

Study ID ☐☐-☐☐☐-2-☐

Consent Form

Please read the information,
Initial the relevant boxes and sign overleaf.

Please sign your initials in boxes

I have read and understood the enclosed Participant Information Sheet version 1.5 1/12/2016. I have had the opportunity to consider the information and have had my questions answered.	☐
I understand that if participate in this study I will not be identified by name in any published reports or papers.	☐
I understand that I am free to withdraw from taking part at any time, without giving a reason.	☐
I understand that my personal details (address) will be used to contact me by post within 3-6 months to complete a follow-up questionnaire.	☐
I understand that sections of my medical notes may be looked at by responsible individuals as part of the study and information may be passed to the research team. I give permission for these individuals to have access to my records. I also understand that sections of my medical records may be looked at as part of study monitoring, or from the ethics committee, or regulatory authorities where it is relevant to my taking part in this research.	☐
I understand that my personal details (date of birth, postcode and NHS number) will be held and stored by the research team at London School of Hygiene & Tropical Medicine and will be supplied to NHS Digital (Previously the Health and Social care Information Centre) to check vital status before I am sent a follow-up questionnaire.	☐
I understand that my personal details (date of birth and NHS number) will be used by the study to link to information that is routinely collected by national clinical audits.	☐
I understand that my questionnaire will be given a unique number that will also be on my consent form, and may be used to link my data.	☐
I understand that the research team will not otherwise release my personal details, except for the above reasons, unless required by law. I will be told if any disclosure was to take place under an exceptional event.	☐

If you have agreed to take part in this study please write your name and address in CAPITAL LETTERS below.

Title _____

First Name _____

Surname _____

Address _____

_____

_____

Postcode _____

Date of birth ___/__/____

Signature _____

Today's date ___/__/2017

FOR COMPLETION BY STAFF:

Affix Patient Label with NHS number in this space

Please ensure patient label contains *their NHS number and Address*

Date of Admission ___/___/ 2017

Name of staff taking consent (in capital letters)

_____

Signature _____

Today's date ___/__/2017

Please now tear out this page

:

Study ID ☐☐-☐☐☐-2-☐

Patients' views of the outcome of care

Cardiology

Inpatient Questionnaire

Q1. What is your date of birth?

D	D	M	M	Y	Y	Y	Y

Please ensure this is your **date of birth NOT** today's date

Q2. Is anyone helping you fill in this questionnaire?

Yes ☐ 1 No ☐ 2

If the answer is yes, please give the relationship to you of the person assisting you

Q3. Are you?

Male ☐ 1 Female ☐ 2

Q4. Which statement best describes your living arrangements?

I live with partner/ spouse/ family/ friends ☐ 1

I live alone ☐ 2

I live in a nursing home, hospital or other long-term care home ☐ 3

Other ☐ 4

Q5 For how long have you had problems with your heart before your current admission?

Less than two weeks	Two weeks to One month	1 month to 12 months	More than 1 year
☐ 1	☐ 2	☐ 3	☐ 4

Q6 Have you had a previous heart attack for which you were admitted to hospital?

Yes ☐ 1 No ☐ 2

We are interested in finding out how you were before your heart attack. Please answer the questions according to how you were before you came into hospital rather than how you are feeling now.

Please answer every question.

Q7. The following is a list of activities that people often do during the week. Although for some people with several medical problems it is difficult to determine what it is that limits them, please go over the activities listed below and how much limitation you have had **due to chest pain, chest tightness or anginal attacks over the past four weeks.**

Place an X in one box on each line.

Activity	Extremely limited	Quite a bit limited	Moderately limited	Slightly limited	Not at all limited	Limited for other reasons or did not do the activity
a. Walking indoors on level ground	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6
b. Gardening, vacuuming or carrying groceries	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6
c. Lifting or moving heavy objects such as. furniture, or lifting children	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6

Q8. Over the past 4 weeks, on average, how many times have you had **chest pain, chest tightness or anginal attacks?**

I have had **chest pain, chest tightness or anginal attacks…**

4 or more times per day	1-3 times per day	3 or more times per week but not every day	1-2 times per week	Less than once a week	None over the past 4 weeks
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6

Q9. Over the past 4 weeks, on average, how many times have you had to take GTN (nitroglycerin tablets or spray) **for your chest pain, chest tightness or anginal attacks?**

I have taken GTN…

4 or more times per day	1-3 times per day	3 or more times per week but not every day	1-2 times per week	Less than once a week	None over the past 4 weeks
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6

Q10. Over the past 4 weeks, how much has **your chest pain, chest tightness or anginal attacks** limited your enjoyment of life?

It has **extremely** limited my enjoyment of life	It has limited my enjoyment of life **quite a bit**	It has **moderately** limited my enjoyment of life	It has **slightly** limited my enjoyment of life	It has **not** limited my enjoyment of life at all
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Remember, we are interested in how you would answer the question below **before your heart attack** rather than how it is right now.

Q11. If you had to spend the rest of your life with your **chest pain, chest tightness or anginal attacks** the way it is right now, how would you feel about this?

Not satisfied at all	Mostly dissatisfied	Somewhat satisfied	Mostly satisfied	Completely satisfied
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

We would like you to think back to <u>four weeks ago</u>, by placing a tick in one box in each group below, please indicate which statements best describe your own health <u>at that time.</u>

12.Mobility

I have no problems in walking about	☐ 1
I have some problems in walking about	☐ 2
I am confined to bed	☐ 3

13.Self-Care

I have no problems with self-care	☐ 1
I have some problems washing or dressing myself	☐ 2
I am unable to wash or dress myself	☐ 3

14.Usual Activities *(e.g. work, study, housework, family or leisure activities)*

I have no problems with performing my usual activities	☐ 1
I have some problems with performing my usual activities	☐ 2
I am unable to perform my usual activities	☐ 3

15.Pain/Discomfort

I have no pain or discomfort	☐ 1
I have moderate pain or discomfort	☐ 2
I have extreme pain or discomfort	☐ 3

16.Anxiety/Depression

I am not anxious or depressed	☐ 1
I am moderately anxious or depressed	☐ 2
I am extremely anxious or depressed	☐ 3

Study ID ☐☐-☐☐☐-2-☐

17. Today's Date

D D M M 2 0 1 7

Please ensure this **is today's date**
NOT your date of birth

18. Have you been told by a doctor that you have any of the following? (Tick all that apply)

Heart disease (for example angina heart attack or heart failure)	☐1	Kidney disease	☐7
High blood pressure	☐2	Diseases of the nervous system (for example Parkinson's disease or multiple sclerosis)	☐8
Problems caused by stroke	☐3	Liver disease	☐9
Leg pain when walking due to poor circulation	☐4	Cancer (within the last 5 years)	☐10
Lung disease (for example asthma, chronic bronchitis or emphysema)	☐5	Depression	☐11
Diabetes	☐6	Arthritis	☐12

You have now finished the questionnaire.
Thank you for your help.

<u>Please return this booklet to the staff who invited you to this study before you leave the hospital.</u>

If you have found this questionnaire, please post this back to:

Chief Investigator
Dr Esther Kwong
London School of Hygiene and Tropical Medicine
15-17 Tavistock Place
London
WC1H 9SH

9.5 Follow-up PROMs Questionnaire for Emergency Laparotomy

Study ID ☐☐-☐☐☐-☐-☐

Version 1.3 3/11/16

Patients' views of the outcome of care

Emergency Laparotomy

Follow-up Questionnaire

Project team:

Dr Esther Kwong (Chief Investigator and project lead)
Professor Nick Black (PhD Supervisor)

London School of Hygiene and Tropical Medicine (LSHTM)

About three months ago you had a hospital admission because of your abdomen (tummy). You are part of a study that asked you how you were before you became unwell at the time; you agreed that we could send you a follow-up questionnaire.

Please can you fill in this questionnaire and return it using the pre-paid envelope. Thank you for your help.

Q1. What is your date of birth?

D	D	M	M	Y	Y	Y	Y

Please ensure this is your **date of birth NOT** today's date

Q2. Is anyone helping you fill in this questionnaire?

Yes ☐ 1 No ☐ 2

If the answer is yes, please give the relationship to you of the person assisting you

Q4. Which statement best describes your living arrangements?

I live with partner/ spouse/ family/ friends ☐ 1

I live alone ☐ 2

I live in a nursing home, hospital or other long-term care home ☐ 3

Other ☐ 4

Q5 Please confirm the date you were admitted to hospital with an operation to your abdomen (tummy) three months ago.

D	D	M	M	2	0	1	7

Please let us know how you are now. If you are unsure about how to answer a question, please give the best answer you can.

Tick <u>one</u> box for <u>every</u> question.

Q6 Have you been readmitted to hospital since the operation on your abdomen (tummy) three months ago?

Yes	No
☐ 1	☐ 2

Q7 Have you had another operation on your abdomen (tummy) since your admission three months ago?

Yes	No
☐ 1	☐ 2

Q8 In general, would you say your health is:

Excellent	Very good	Good	Fair	Poor
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q9 How would you describe the result of operation from your admission three months ago?

Excellent	Very good	Good	Fair	Poor
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q10. Overall, how are the problems now with your abdomen (tummy), compared to one month before your admission?

Excellent	Very good	Good	Fair	Poor
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q11. How often during the past 2 weeks have you had pain in the abdomen?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q12. How often during the past 2 weeks have you had a feeling of fullness in the abdomen?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q13. How often during the past 2 weeks have you had bloating (sensation of too much gas in the abdomen)?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q14. How often during the past 2 weeks have you been troubled by excessive passage of gas through the anus?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q15. How often during the past 2 weeks have you been troubled by strong burping or belching?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q.16. How often during the past 2 weeks have you been troubled by gurgling noises from the abdomen?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q17. How often during the past 2 weeks have you been troubled by frequent bowel movements?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q18. How often during the past 2 weeks have you found eating to be a pleasure?

Never	Little of the time	Some of the time	Most of the time	All of the time
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q19. Because of your illness, to what extent have you restricted the kinds of food you eat?

Very much	Much	Somewhat	Little	Not at all
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q20. During the past 2 weeks, how well have you been able to cope with everyday stresses?

Extremely Poorly	Poorly	Moderately	Well	Extremely Well
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q21. How often during the past 2 weeks have you been sad about being ill?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q22. How often during the past 2 weeks have you been nervous or anxious about your illness?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q23. How often during the past 2 weeks have you been happy with life in general?

Never	Little of the time	Some of the time	Most of the time	All of the time
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q24. How often during the past 2 weeks have you been frustrated about your illness?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q25. How often during the past 2 weeks have you been tired or fatigued?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q26. How often during the past 2 weeks have you felt unwell?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q27. Over the past week, have you woken up in the night?

Every night	5-6 nights	3-4 nights	1-2 nights	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q28. Since becoming ill, have you been troubled by changes in your appearance?

A great deal	A moderate amount	Somewhat	A little bit	Not at all
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q29. Because of your illness, how much physical strength have you lost?

A great deal	A moderate amount	Somewhat	A little bit	Not at all
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q30. Because of your illness, to what extent have you lost your endurance?

A great deal	A moderate amount	Somewhat	A little bit	Not at all
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q31. Because of your illness, to what extent do you feel unfit?

Extremely Unfit	Unfit	Moderately Unfit	A little Unfit	Fit
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q32. During the past 2 weeks, how often have you been able to complete your normal daily activities (school, work, household)?

Never	Little of the time	Some of the time	Most of the time	All of the time
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q33. During the past 2 weeks, how often have you been able to take part in your usual patterns of leisure or recreational activities?

Never	Little of the time	Some of the time	Most of the time	All of the time
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q34. During the past 2 weeks, how much have you been troubled by the medical treatment of your illness?

Very much	Much	Somewhat	Little	Not at all
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q35. To what extent have your personal relations with people close to you (family or friends) worsened because of your illness?

Very much	Much	Somewhat	Little	Not at all
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q36. To what extent has your sexual life been impaired (harmed) because of your illness?

Very much	Much	Somewhat	Little	Not at all
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q37. How often during the past 2 week, have you been troubled by fluid or food coming up into your mouth (regurgitation)?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q38. How often during the past 2 weeks have you felt uncomfortable because of your slow speed of eating?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q39. How often during the past 2 weeks have you had trouble swallowing your food?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q40. How often during the past 2 weeks have you been troubled by urgent bowel movements?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q41. How often during the past 2 weeks have you been troubled by diarrhoea?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q42. How often during the past 2 weeks have you been troubled by constipation?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q43. How often during the past 2 weeks have you been troubled by nausea?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q44. How often during the past 2 weeks have you been troubled by blood in the stool?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q45. How often during the past 2 weeks have you been troubled by heartburn?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Q46. How often during the past 2 weeks have you been troubled by uncontrolled stools?

All of the time	Most of the time	Some of the time	Little of the time	Never
☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Please continue to the next page

By placing a tick in one box in each group below, please indicate which statements best describe your own health state today.

47.Mobility

I have no problems in walking about	☐ 1
I have some problems in walking about	☐ 2
I am confined to bed	☐ 3

48.Self-Care

I have no problems with self-care	☐ 1
I have some problems washing or dressing myself	☐ 2
I am unable to wash or dress myself	☐ 3

49.Usual Activities *(e.g. work, study, housework, family or leisure activities)*

I have no problems with performing my usual activities	☐ 1
I have some problems with performing my usual activities	☐ 2
I am unable to perform my usual activities	☐ 3

50.Pain/Discomfort

I have no pain or discomfort	☐ 1
I have moderate pain or discomfort	☐ 2
I have extreme pain or discomfort	☐ 3

51.Anxiety/Depression

I am not anxious or depressed	☐ 1
I am moderately anxious or depressed	☐ 2
I am extremely anxious or depressed	☐ 3

Study ID ☐☐-☐☐☐-☐-☐

52. Today's Date

D D M M 2 0 Y Y

Please ensure this **is today's date**
NOT your date of birth

You have now finished the questionnaire.

Thank you for your help.

Please return this questionnaire in the envelope provided. You do not have to use a stamp, the postage is already paid.

Contact us for further details

[Insert contact details for chief investigator]

9.6 Follow-up PROMs Questionnaire for STEMI

Study ID 2 3 - 0 8 1 - 2 - 3

Version 1.4 3/11/16

Patients' views of the outcome of care

Cardiology

Follow-up Questionnaire

Project team:

Dr Esther Kwong (Chief Investigator and project lead)
Professor Nick Black (PhD Supervisor)

London School of Hygiene and Tropical Medicine (LSHTM)

About three months ago you had a hospital admission because of your heart. You are part of a study that asked you how you were before you became unwell at the time; you agreed that we could send you a follow-up questionnaire.

Please can you fill in this questionnaire and return it using the pre-paid envelope. Thank you for your help.

Q1. What is your date of birth?

D	D	M	M	Y	Y	Y	Y

Please ensure this is your **date of birth NOT** today's date

Q2. Is anyone helping you fill in this questionnaire?

Yes ☐ 1 No ☐ 2

If the answer is yes, please give the relationship to you of the person assisting you

Q4. Which statement best describes your living arrangements?

I live with partner/ spouse/ family/ friends ☐ 1

I live alone ☐ 2

I live in a nursing home, hospital or other long-term care home ☐ 3

Other ☐ 4

Q5 Please confirm the date you were admitted to hospital with a heart attack three months ago.

D	D	M	M	2	0	1	7

Please let us know how you are now. If you are unsure about how to answer a question, please give the best answer you can.

Tick <u>one</u> box for <u>every</u> question.

Q6 Have you been readmitted to hospital since your heart attack three months ago?

Yes ☐ 1 No ☐ 2

Q7 Have you had another heart attack since your admission three months ago?

Yes ☐ 1 No ☐ 2

Q8 In general, would you say your health is:

Excellent ☐ 1 Very good ☐ 2 Good ☐ 3 Fair ☐ 4 Poor ☐ 5

Q9 How would you describe the result of treatment from your admission three months ago?

Excellent ☐ 1 Very good ☐ 2 Good ☐ 3 Fair ☐ 4 Poor ☐ 5

Q10 Overall, how are the problems now with your heart, compared to one month before you had the heart attack?

Excellent	Very good	Good	Fair	Poor
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q11. The following is a list of activities that people often do during the week. Although for some people with several medical problems it is difficult to determine what it is that limits them, please go over the activities listed below and how much limitation you have had **due to chest pain, chest tightness or angina over the past four weeks.**

Activity	Extremely limited	Quite a bit limited	Moderately limited	Slightly limited	Not at all limited	Limited for other reasons or did not do the activity
a. Walking indoors on level ground	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 9
b. Gardening, vacuuming or carrying groceries	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 9
c. Lifting or moving heavy objects (e.g. furniture, children)	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 9

Q12. Over the past 4 weeks, on average, how many times have you had **chest pain, chest tightness or angina?**

I have had **chest pain, chest tightness or angina…**

4 or more times per day	1-3 times per day	3 or more times per week but not every day	1-2 times per week	Less than once a week	None over the past 4 weeks
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6

Q13. Over the past 4 weeks, on average, how many times have you had to take glyceryl trinitrate (GTN tablets or spray) **for your chest pain, chest tightness or angina?**

I have taken nitroglycerin…

4 or more times per day	1-3 times per day	3 or more times per week but not every day	1-2 times per week	Less than once a week	None over the past 4 weeks
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6

Q14. Over the past 4 weeks, how much has **your chest pain, chest tightness or angina** limited your enjoyment of life?

It has **extremely** limited my enjoyment of life	It has limited my enjoyment of life **quite a bit**	It has **moderately** limited my enjoyment of life	It has **slightly** limited my enjoyment of life	It has **not** limited my enjoyment of life at all
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q15. If you had to spend the rest of your life with your **chest pain, chest tightness or angina** the way it is right now, how would you feel about this?

Not satisfied at all	Mostly dissatisfied	Somewhat satisfied	Mostly satisfied	Completely satisfied
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

By placing a tick in one box in each group below, please indicate which statements best describe your own health state today.

16.Mobility

I have no problems in walking about ☐1

I have some problems in walking about ☐2

I am confined to bed ☐3

17.Self-Care

I have no problems with self-care ☐1

I have some problems washing or dressing myself ☐2

I am unable to wash or dress myself ☐3

18.Usual Activities *(e.g. work, study, housework, family or leisure activities)*

I have no problems with performing my usual activities ☐1

I have some problems with performing my usual activities ☐2

I am unable to perform my usual activities ☐3

19.Pain/Discomfort

I have no pain or discomfort ☐1

I have moderate pain or discomfort ☐2

I have extreme pain or discomfort ☐3

20.Anxiety/Depression

I am not anxious or depressed ☐1

I am moderately anxious or depressed ☐2

I am extremely anxious or depressed ☐3

Study ID 2 3 - 0 8 1 - 2 - 3

21. Today's Date

D D M M 2 0 Y Y

Please ensure this **is today's date**
NOT your date of birth

You have now finished the questionnaire.

Thank you for your help.

Please return this questionnaire in the envelope provided. You do not have to use a stamp, the postage is already paid.

Contact us for further details:

Chief Investigator
Dr Esther Kwong
London School of Hygiene and Tropical Medicine
15-17 Tavistock Place
London
WC1H 9SH

Email:esther.kwong@lshtm.ac.uk
T:020 7985 8285

9.7 Appendices relevant to Chapter 4

9.7.1 Tables

Tables A-D

Table A: Retrospective PROMs compared with GPPS data (England)

Group England Hip	Number in group - questionnaire	Number in group - GPPS data	EQ5D Mean (SD) - QR	Mean (SD) - GPPS data	Difference in means (95% CI)
Men, 60 or under	22	83975	0.31 (0.35)	0.79 (0.28)	0.48 (0.33 - 0.63)
Men, 61-75	30	81906	0.26 (0.37)	0.79 (0.20)	0.53 (0.40 - 0.67)
Men, 76 and above	23	39361	0.34 (0.36)	0.71 (0.24)	0.38 (0.22 - 0.53)
Women, 60 or under	32	100277	0.22 (0.37)	0.74 (0.27)	0.51 (0.38 - 0.65)
Women, 61-75	67	211827	0.18 (0.33)	0.75 (0.24)	0.56 (0.48 - 0.64)
Women, 76 and above	53	71711	0.16 (0.33)	0.65 (0.23)	0.48 (0.40 - 0.58)

* 95% CI calculated using diff +/- 1.96*SE(diff). SE(diff) = sqrt(SD_q1^2/n1 + SD_q2^2/n2)

Table B: Contemporary PROMs compared with GPPS data (England)

Group England Hip	Number in group - questionnaire	Number in group - GPPS data	EQ5D Mean (SD) – Q1	Mean (SD) - GPPS data	Difference in means (95% CI)
Men, 60 or under	21	83975	0.39 (0.37)	0.79 (0.28)	0.40 (0.23 - 0.56)
Men, 61-75	26	81906	0.32 (0.34)	0.79 (0.22)	0.47 (0.33 - 0.61)
Men, 76 and above	25	39361	0.18 (0.32)	0.71(0.24)	0.53 (0.40 - 0.67)
Women, 60 or under	28	100277	0.25 (0.36)	0.74 (0.29)	0.48 (0.34 - 0.62)
Women, 61-75	63	211827	0.23 (0.31)	0.75 (0.24)	0.52 (0.43 - 0.59)
Women, 76 and above	48	71711	0.14 (0.32)	0.65 (0.25)	0.50 (0.41 -0.60)

Table C: Retrospective PROMs compared with GPPS data (matched for local authority)

Group local Hip	Number in group – PROMs	Number in group - GPPS data	EQ5D Mean (SD) - QR	Mean (SD) - GPPS data	Difference in means (95% CI)
Men, 60 or under	22	10789	0.317 (0.35)	0.79 (0.27)	0.48 (0.33 - 0.63)
Men, 61-75	31	8151	0.26 (0.37)	0.80 (0.20)	0.55 (0.41 - 0.68)
Men, 76 and above	23	3324	0.34 (0.36)	0.71 (0.25)	0.48 (0.32 - 0.63)
Women, 60 or under	31	13644	0.23 (0.37)	0.74 (0.28)	0.51 (0.37 - 0.64)
Women, 61-75	67	20778	0.18 (0.33)	0.74 (0.24)	0.55 (0.47 - 0.64)
Women, 76 and above	54	7313	0.15 (0.33)	0.63 (0.25)	0.48 (0.39 - 0.57)

Table D: Contemporary PROMs compared with GPPS data (matched for local authority)

Group local Hip	Number in group - questionnaire	Number in group - GPPS data	EQ5D Mean (SD) – Q1	Mean (SD) - GPPS data	Difference in means (95% CI)
Men, 60 or under	21	10789	0.39 (0.37)	0.79 (0.28)	0.40 (0.23 - 0.56)
Men, 61-75	26	8151	0.32 (0.34)	0.80 (0.22)	0.48 (0.34 - 0.62)
Men, 76 and above	25	3324	0.18 (0.32)	0.71 (0.25)	0.53 (0.40 - 0.67)
Women, 60 or under	27	13644	0.27 (0.35)	0.74 (0.30)	0.46 (0.35 - 0.58)
Women, 61-75	63	20778	0.23 (0.31)	0.74 (0.25)	0.51 (0.43 - 0.59)
Women, 76 and above	49	7313	0.13 (0.32)	0.63 (0.26)	0.50 (0.40 - 0.59)

Tables E-H: Matched by Local area VS England wide (by gender and SES)

Table E: Retrospective PROMs compared with GPPS data (England)

Group England Hip	Number in group - questionnaire	Number in group - GPPS data	EQ5D Mean (SD) - QR	Mean (SD) - GPPS data	Difference in means (95% CI)
Men, 1 (least deprived)	9	21520	0.30 (0.38)	0.77 (0.23)	0.47 (0.18 - 0.77)
Women, 1 (least deprived)	10	26013	0.19 (0.30)	0.80 (0.18)	0.60 (0.40 - 0.81)
Men, 2	8	28884	0.37 (0.37)	0.82 (0.21)	0.49 (0.18 - 0.80)
Women, 2	30	73867	0.17 0.37)	0.74 (0.23)	0.58 (0.44 0.72)
Men, 3	25	78031	0.32 (0.39)	0.78 (0.24)	0.47 (0.30 - 0.63)
Women, 3	40	114489	0.19 (0.34)	0.75 (0.23)	0.56 (0.45 - 0.67)
Men, 4	16	34868	0.40 (0.30)	0.79 (0.20)	0.39 (0.24 - 0. 55)
Women, 4	42	100775	0.23 (0.33)	0.71 (0.26)	0.47 (0.37 - 0.58)
Men, 5 (most deprived)	17	41894	0.13 (0.33)	0.70 (0.28)	0.56 (0.40 - 0.73)
Women, 5 (most deprived)	29	68671	0.11 (0.33)	0.67 (0.26)	0.55 (0.43 - 0.68)

Table F: Contemporary PROMs compared with GPPS data (England)

Group England Hip	Number in group - questionnaire	Number in group - GPPS data	EQ5D Mean (SD) – Q1	Mean (SD) - GPPS data	Difference in means (95% CI)
Men, 1 (least deprived)	10	21520	0.38 (0.30)	0.77 (0.23)	0.70 (0.48 - 0.91)
Women, 1 (least deprived)	7	26013	0.35 (0.34)	0.80 (0.22)	0.45 (0.13 - 0.77)
Men, 2	8	28884	0.33 (0.37)	0.82 (0.21)	0.49 (0.18 - 0.81)
Women, 2	28	73867	0.22 (0.38)	0.74 (0.24)	0.52 (0.38 - 0 .67)
Men, 3	25	78031	0.33 (0.35)	0.78 (0.25)	0.45 (0 .31 - 0.60)
Women, 3	37	114489	0.17 (0.35)	0.75 (0.24)	0.58 (0.46 - 0.70)
Men, 4	13	34868	0.46 (0.29)	0.79 (0.22)	0.32 (0.15 - 0.51)
Women, 4	38	100775	0.24 (0.29)	0.71 (0.27)	0.46 (0.37 - 0.56)
Men, 5 (most deprived)	16	41894	0.21 (0.36)	0.70 (0.29)	0.48 (0.29 - 0.68)
Women, 5 (most deprived)	29	68671	0.14 (0.29)	0.67 (0.27)	0.53 (0.41 - 0.64)

Table G: Retrospective PROMs compared with GPPS data (matched for local authority)

Group local Hip	Number in group - questionnaire	Number in group - GPPS data	EQ5D Mean (SD) - QR	Mean (SD) - GPPS data	Difference in means (95% CI)
Men, 1 (least deprived)	9	1347	0.30 (0.38)	0.77 (0.22)	0.47 (0.18 - 0.77)
Women, 1 (least deprived)	10	1460	0.19 (0.29)	0.80 (0.20)	0.60 (0.40 - 0.81)
Men, 2	8	2299	0.37 (0.37)	0.82 (0.21)	0.49 (0.18 - 0.80)
Women, 2	30	5812	0.17 (0.37)	0.73 (0.24)	0.57 (0.43 - 0.71)
Men, 3	25	7233	0.32 (0.39)	0.80 (0.25)	0.48 (0.32 - 0.64)
Women, 3	40	10387	0.19 (0.34)	0.74 (0.24)	0.55 (0.44 - 0.66)
Men, 4	16	4997	0.40 (0.30)	0.80 (0.20)	0.41 (0.26 - 0.57)
Women, 4	42	14126	0.23 (0.33)	0.71 (0.27)	0.47 (0.37 - 0.58)
Men, 5 (most deprived)	17	6388	0.13 (0.33)	0.73 (0.27)	0.60 (0.43 - 0.76)
Women, 5 (most deprived)	29	9950	0.11 (0.33)	0.69 (0.27)	0.57 (0.45 - 0.70)

Table H: Contemporary PROMs compared with GPPS data (matched for local authority)

Group local Hip	Number in group - questionnaire	Number in group - GPPS data	EQ5D Mean (SD) – Q1	Mean (SD) - GPPS data	Difference in means (95% CI)
Men, 1 (least deprived)	10	1347	0.30 (0.30)	0.77 (0.22)	0.47 (0.48 - 0.77)
Women, 1 (least deprived)	7	1460	0.35 (0.35)	0.80 (0.24)	0.44 (0.13 - 0.77)
Men, 2	8	2299	0.33 (0.37)	0.82 (0.21)	0.49 (0.18 - 0.81)
Women, 2	28	5812	0.22 (0.38)	0.73 (0.25)	0.51 (0.37 - 0.66)
Men, 3	25	7233	0.33 (0.35)	0.80 (0.25)	0.47 (0.32 - 0.61)
Women, 3	37	10387	0.17 (0.35)	0.74(0.25)	0.57 (0.46 - 0.68)
Men, 4	13	4997	0.46 (0.29)	0.81 (0.22)	0.34 (0.19 - 0.50)
Women, 4	38	14126	0.24 (0.29)	0.71 (0.28)	0.46 (0.38 - 0.56)
Men, 5 (most deprived)	16	6388	0.21 (0.36)	0.73 (0.29)	0.52 (0.34 - 0.71)
Women, 5 (most deprived)	29	9950	0.14 (0.29)	0.69 (0.27)	0.55 (0.44 - 0.66)

Tables I-P: Handling of co-morbidities by count and by exact methods

Table I: Co-morbidities by count retrospective PROMs with GPPS (England, Method 1 treatment for primary condition)

Group Count 1 England Hip	Number in group - questionnaire	Number in group - GPPS data	EQ5D Mean (SD) - QR	Mean (SD) - GPPS data	Difference in means (95% CI)
Men, 60 or under	22	83975	0.31 (0.35)	0.79 (0.28)	0.48 (0.33 - 0.63)
Men, 61-75	30	81906	0.26 (0.37)	0.79 (0.20)	0.53 (0.40 - 0.67)
Men, 76 and above	23	39361	0.34 (0.36)	0.71 (0.25)	0.38 (0.22 - 0.53
Women, 60 or under	32	100277	0.22 (0.37)	0.74 (0.27)	0.51 (0.38 - 0.65)
Women, 61-75	67	211827	0.18 (0.33)	0.75 (0.24)	0.56 (0.48 - 0.65)
Women, 76 and above	53	71711	0.16 (0.33)	0.65 (0.23)	0.48 (0.40 - 0.58)

Table J: Co-morbidities by count contemporary PROMs with GPPS (England, Method 1 treatment for primary condition)

Group Count 1 England Hip	Number in group - questionnaire	Number in group - GPPS data	EQ5D Mean (SD) – Q1	Mean (SD) - GPPS data	Differences in means (95% CI)
Men, 60 or under	21	83975	0.39 (0.37)	0.79 (0.28)	0.40 (0.23 - 0.56)
Men, 61-75	26	81906	0.32 (0.34)	0.79 (0.22)	0.47 (0.33 - 0.61)
Men, 76 and above	25	39361	0.18 (0.32)	0.71 (0.24)	0.53 (0.40 - 0.67)
Women, 60 or under	28	100277	0.25 (0.36)	0.74 (0.29)	0.48 (0.34 - 0.62)
Women, 61-75	63	211827	0.23 (0.31)	0.75 (0.24)	0.52 (0.43 - 0.60)
Women, 76 and above	48	71711	0.14 (0.32)	0.65 (0.25)	0.50 (0.41 - 0.60)

Table K: Co-morbidities by exact match, retrospective PROMs with GPPS (England, Method 1 treatment for primary condition)

Group Exact 1 England Hip	Number in group - questionnaire	Number in group - GPPS data	EQ5D Mean (SD) - QR	Mean (SD) - GPPS data	Differences in means (95% CI)
Men, 60 or under	23	33307	0.32 (0.35)	0.85 (0.25)	0.53 (0.38 - 0.68)
Men, 61-75	32	22879	0.24 (0.37)	0.83 (0.21)	0.59 (0.45 - 0.72)
Men, 76 and above	24	7430	0.32 (0.36)	0.77 (0.22)	0.45(0.29 - 0.60)
Women, 60 or under	33	31122	0.23 (0.37)	0.79 (0.27)	0.56 (0.43 - 0.69)
Women, 61-75	72	64308	0.20 (0.34)	0.74 (0.23)	0.54 (0.47 - 0.63)
Women, 76 and above	56	19645	0.17 (0.33)	0.65 (0.23)	0.48 (0.40 - 0.57)

Table L: Co-morbidities by exact match, contemporary PROMs with GPPS (England, Method 1 treatment for primary condition)

Group Exact 1 England Hip	Number in group - questionnaire	Number in group - GPPS data	EQ5D Mean (SD) – Q1	Mean (SD) - GPPS data	Differences in means (95% CI)
Men, 60 or under	21	33307	0.3909 (0.3672)	0.85 (0.25)	0.46 (0.31 - 0.61)
Men, 61-75	26	22879	0.3213(0.3430)	0.83 (0.21)	0.50 (0.37 - 0.64)
Men, 76 and above	25	7430	0.1783 (0.3199)	0.77 (0.22)	0.59 (0.46 - 0.72)
Women, 60 or under	28	31122	0.2541 (0.3588)	0.79 (0.27)	0.53 (0.28 - 0.80)
Women, 61-75	63	64308	0.2253 (0.3118)	0.74 (0.23)	0.51 (0.44 - 0.60)
Women, 76 and above	48	19645	0.1426 (0.3229)	0.65 (0.23)	0.50 (0.41 - 0.60)

Table M: Co-morbidities by count, contemporary PROMs with GPPS (England, Method 3 treatment for primary condition)

Group Count 3 England Hip	Number in group - questionnaire	Number in group - GPPS data	EQ5D Mean (SD) - QR	Mean (SD) - GPPS data	Differences in means (95% CI)
Men, 60 or under	23	131879	0.3244 (0.3467)	0.84 (0.24)	0.52 (0.37 -0.67)
Men, 61-75	32	96312	0.2401 (0.3679)	0.80 (0.20)	0.56 (0.43 - 0.70)
Men, 76 and above	24	53934	0.3237 (0.3621)	0.74 (0.24)	0.42(0.26 - 0.57)
Women, 60 or under	33	179328	0.2290 (0.3724)	0.81 (0.24)	0.58 (0.44 - 0.71)
Women, 61-75	72	290283	0.1963 (0.3410)	0.78 (0.23)	0.58 (0.50 - 0.66)
Women, 76 and above	56	101061	0.1650 (0.3289)	0.68 (0.24)	0.58 (0.43 - 0.61)

Table N: Co-morbidities by count, contemporary PROMs with GPPS (England, Method 3 treatment for primary condition)

Group Count 3 England Hip	Number in group - questionnaire	Number in group - GPPS data	EQ5D Mean (SD) – Q1	Mean (SD) - GPPS data	Differences in means (95% CI)
Men, 60 or under	21	131879	0.39 (0.37)	0.84 (0.24)	0.45 (0.28 - 0.62)
Men, 61-75	26	96312	0.32 (0.34)	0.80 (0.20)	0.48 (0.33 - 0.62)
Men, 76 and above	25	53934	0.18 (0.32)	0.74 (0.24)	0.44 (0.32 - 0.58)
Women, 60 or under	28	179328	0.25 (0.36)	0.81 (0.24)	0.56 (0.43 - 0.70)
Women, 61-75	63	290283	0.23 (0.34)	0.77 (0.23)	0.55 (0.42 - 0.67)
Women, 76 and above	48	101061	0.14 (0.32)	0.68 (0.24)	0.54 (0.45 - 0.63)

Table O: Co-morbidities by exact match, retrospective PROMs with GPPS (England, Method 3 treatment for primary condition)

Group Exact 3 England Hip	Number in group - questionnaire	Number in group - GPPS data	EQ5D Mean (SD) - QR	Mean (SD) - GPPS data	Differences in means (95% CI)
Men, 60 or under	23	104573	0.37 (0.37)	0.87 (0.22)	0.50 (0.34 - 0.66)
Men, 61-75	33	57554	0.30 (0.34)	0.85 (0.20)	0.55 (0.43 - 0.67)
Men, 76 and above	26	17056	0.17 (0.32)	0.78 (0.23)	0.61(0.48 - 0.74)
Women, 60 or under	33	124128	0.25 (0.36)	0.85 (0.24)	0.61 (0.48 - 0.73)
Women, 61-75	72	187045	0.21 (0.31)	0.82 (0.22)	0.61 (0.54 - 0.69)
Women, 76 and above	57	48805	0.13 (0.34)	0.74 (0.23)	0.60 (0.51 - 0.70)

Table P: Co-morbidities by exact match, contemporary PROMs with GPPS (England, Method 3 treatment for primary condition)

Group Exact 3 England Hip	Number in group - questionnaire	Number in group - GPPS data	EQ5D Mean (SD) – Q1	Mean (SD) - GPPS data	Differences in means (95% CI)
Men, 60 or under	23	104573	0.32 (0.35)	0.87 (0.22)	0.55 (0.40 - 0.69)
Men, 61-75	33	57554	0.24 (0.37)	0.85 (0.20)	0.61 (0.48 - 0.74)
Men, 76 and above	24	17056	0.32 (0.36)	0.78 (0.23)	0.44 (0.32 - 0.58)
Women, 60 or under	33	124128	0.23 (0.37)	0.85 (0.24)	0.46 (0.31 - 0.61)
Women, 61-75	72	187045	0.20 (0.34)	0.82 (0.22)	0.62 (0.49 - 0.75)
Women, 76 and above	56	48805	0.17 (0.33)	0.74 (0.23)	0.63 (0.54 - 0.71)

Tables Q-T: Method 2 of handling of primary condition (Treat all patients in PROMs cohort as having arthritis; mapping to GPPS patients with arthritis

Table Q: Primary condition arthritis matched retrospective PROMs with GPPS (England, other co-morbidities handled by count)

Group Count 2 England Hip	Number in group - questionnaire	Number in group - GPPS data	EQ5D Mean (SD) - QR	Mean (SD) - GPPS data	Differences in means (95% CI)
Men, 60 or under	23	13121	0.32 (0.35)	0.59 (0.29)	0.26 (0.11 - 0.41)
Men, 61-75	32	16866	0.24 (0.37)	0.59 (0.23)	0.35 (0.22 - 0.48)
Men, 76 and above	24	13671	0.32 (0.36)	0.59 (0.25)	0.26 (0.11 - 0.41)
Women, 60 or under	33	25128	0.23 (0.37)	0.55 (0.29)	0.32 (0.19 - 0.45)
Women, 61-75	72	83677	0.20 (0.34)	0.60 (0.25)	0.41 (0.33 - 0.49)
Women, 76 and above	56	40937	0.17 (0.33)	0.56 (0.24)	0.39(0.30 - 0.48)

Table R: Primary condition arthritis matched contemporary PROMs with GPPS (England, other co-morbidities handled by count)

Group Count 2 England Hip	Number in group - questionnaire	Number in group - GPPS data	EQ5D Mean (SD) – Q1	Mean (SD) - GPPS data	Differences in means (95% CI)
Men, 60 or under	22	13121	0.37 (0.37)	0.59 (0.29)	0.21 (0.05 - 0.38)
Men, 61-75	28	16866	0.30 (0.34)	0.59 (0.22)	0.29 (0.16 - 0.43)
Men, 76 and above	26	13671	0.17 (0.32)	0.59 (0.25)	0.41 (0.29 - 0.54)
Women, 60 or under	29	25128	0.25 (0.36)	0.55 (0.29)	0.31 (0.17 - 0.44)
Women, 61-75	68	83677	0.21 (0.31)	0.60 (0.25)	0.40 (0.32 - 0.47)
Women, 76 and above	51	40937	0.13 (0.32)	0.56 (0.24)	0.42 (0.33 - 0.51)

Table S: Primary condition arthritis matched retrospective PROMs with GPPS (England, other co-morbidities handled by exact match)

Group Exact 2 England Hip	Number in group - questionnaire	Number in group - GPPS data	EQ5D Mean (SD) - QR	Mean (SD) - GPPS data	Differences in means (95% CI)
Men, 60 or under	23	8149	0.32 (0.35)	0.64 (0.27)	0.31 (0.16 - 0.46)
Men, 61-75	32	8762	0.24 (0.37)	0.66 (0.22)	0.41 (0.28 - 0.55)
Men, 76 and above	24	4201	0.32 (0.36)	0.63 (0.24)	0.31(0.16 - 0.45)
Women, 60 or under	33	12743	0.23 (0.37)	0.61 (0.26)	0.37 (0.25 - 0.51)
Women, 61-75	72	48553	0.20 (0.34)	0.67 (0.24)	0.46 (0.39 - 0.55)
Women, 76 and above	56	17949	0.17 (0.33)	0.61 (0.22)	0.44 (0.36 - 0.54)

Table T: Primary condition arthritis matched contemporary PROMs with GPPS (England other co-morbidities handled by exact match)

Group Exact 2 England Hip	Number in group - questionnaire	Number in group - GPPS data	EQ5D Mean (SD) – Q1	Mean (SD) - GPPS data	Differences in means (95% CI)
Men, 60 or under	22	8149	0.39 (0.37)	0.64 (0.27)	0.24 (0.08 - 0.41) p=0.048
Men, 61-75	28	8762	0.32 (0.34)	0.66 (0.22)	0.33 (0.20 - 0.47)
Men, 76 and above	26	4201	0.18 (0.32)	0.63 (0.24)	0.44 (0.32 - 0.58)
Women, 60 or under	29	12743	0.25 (0.36)	0.61 (0.26)	0.35 (0.21 - 0.49)
Women, 61-75	68	48553	0.23 (0.31)	0.67 (0.24)	0.44 (0.36 - 0.52)
Women, 76 and above	51	17949	0.14 (0.32)	0.61 (0.22)	0.46 (0.36 - 0.52)

9.7.2 Further Statistical Analysis Plan for Emergency Patients Cohort groups

Further research Question to answer in emergency cohorts: Can GPPS population means offer a more reliable alternative to individual retrospectively collected measures of patient reported outcomes in emergency admission cohorts.

In order to test for agreement using the ICC (agreement, consistency) between population mean EQ-5D with individual patient reported outcome measures, for the analysis, we will be making two pragmatic assumptions:

Firstly, we are treating retrospective PROMs as though they were a gold standard, even though we know that they are not necessarily a perfect measure. However, in practice, for people with an emergency admission, this is the closest recording of patients' baseline health status from their own perspective we can obtain. We also know from elective patients that the agreement between their contemporary and retrospective PROMs is very strong.

Secondly, we ignore the variance within the matched GPPS groups since the group sample sizes are large. E.g. if the SD was 0.26 and the sample size was 4000 the SE would be 0.004, meaning that this is negligible for the ICC calculation. Consequently, the residual variance in the individual PROMs scores is likely to be much larger than the residual variance in the matched group means, (we know that the residual error between the matched mean GPPS population EQ5D is smaller and also different for each match when compared to the residual error in individual PROMs).

The issue is that the ICC formulae tend to assume one constant error between patients (the rows) and "raters" (in this case, the individual patient themselves vs the matched GPPS group mean). However, the row variation between the matched GPPS means reflects systematic variation related to observed group characteristics, and so does not provide a useful contribution to the estimated of the residual variance.

The form of the intra-class correlation we therefore would use is: (〖MS〗_patients-〖MS〗_error)/(〖MS〗_patients+〖MS〗_error), where the mean squared variation between patients is calculated from the individual level retrospective PROM data only, and the mean squared error is calculated from the squared differences between the individual PROM and matched GPPS mean.